

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 22:05:56 ; Search time 29 Seconds  
(without alignments)  
231.325 Million cell updates/sec

Title: us-09-821-255-2

Perfect score: 1211

Sequence: 1 ILGGREAEHARPYMASVOL.....KPGIYTRVASYAAWIDSVLA 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

al number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2.6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2.6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2.6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2.6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2.6/ptodata/1/iaa/PTUS\_COMB.pep.\*
- 6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1205	99.5	238	6	5223425-5
2	1205	99.5	250	6	5223425-4
3	1201	99.2	253	6	5223425-8
4	1198	98.9	228	4	US-08-944-483-44
5	1051.5	86.8	222	6	5223425-6
6	844.5	69.7	259	6	5223425-2
7	844	69.7	260	6	5223425-10
8	471.5	38.9	232	4	US-08-944-483-45
9	456.5	37.7	234	4	US-08-944-483-46
10	452.5	37.4	232	1	US-07-990-301A-4
11	445	36.7	258	1	US-07-990-301A-2
12	440.5	36.4	260	4	US-09-008-271A-7
13	433.5	35.8	244	1	US-08-361-395-1
14	433.5	35.8	260	3	US-09-070-526-2
15	421	34.8	288	4	US-09-386-642-13
16	419.5	34.6	260	3	US-09-025-059-3
17	416.5	34.4	247	2	US-08-956-267A-2
18	415	34.3	237	3	US-08-768-859A-16
19	415	34.3	237	3	US-08-767-820A-16
20	415	34.3	237	3	US-08-622-046B-12
21	415	34.3	237	4	US-08-944-483-37
22	415	34.3	237	4	US-09-100-264-1
23	415	34.3	238	3	US-08-768-859A-8
24	415	34.3	238	3	US-08-767-820A-8
25	415	34.3	244	3	US-08-768-859A-10
26	415	34.3	244	3	US-08-767-820A-10
27	415	34.3	244	3	US-08-622-046B-16

28	415	34.3	244	4	US-09-100-264-5	Sequence 5, Appli
29	415	34.3	261	3	US-08-768-859A-6	Sequence 6, Appli
30	415	34.3	261	3	US-08-767-820A-6	Sequence 6, Appli
31	415	34.3	261	3	US-08-622-046B-14	Sequence 14, Appli
32	415	34.3	261	4	US-09-100-264-7	Sequence 7, Appli
33	415	34.3	261	4	US-08-983-075D-7	Sequence 7, Appli
34	414.5	34.2	237	5	PCT-US94-07329-11	Sequence 11, Appli
35	414.5	34.2	237	5	US-08-096-946-11	Sequence 11, Appli
36	414.5	34.2	237	5	PCT-US95-06157-1	Sequence 1, Appli
37	414	34.2	237	1	US-08-096-946-10	Sequence 10, Appli
38	414	34.2	237	5	PCT-US94-07329-10	Sequence 10, Appli
39	414	34.2	237	5	PCT-US95-06157-16	Sequence 16, Appli
40	413.5	34.1	237	3	US-08-768-859A-1	Sequence 1, Appli
41	413.5	34.1	237	3	US-08-767-820A-1	Sequence 1, Appli
42	413.5	34.1	237	3	US-08-622-046B-7	Sequence 7, Appli
43	413.5	34.1	237	4	US-08-944-483-38	Sequence 38, Appli
44	413.5	34.1	237	4	US-09-100-264-3	Sequence 3, Appli
45	413.5	34.1	237	4	US-09-303-339-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
5223425-5  
; Patent No. 5223425  
; APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,  
; BARRY M.; WHITE, TYLER R.  
; TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT  
; D ACTIVITY  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/277,963  
; FILING DATE: 30-NOV-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 34,203  
; FILING DATE: 02-APR-1987  
; SEQ ID NO: 5:  
; LENGTH: 238  
5223425-5

Query Match	99.5%	Score 1205;	DB 6;	Length 238;
Best Local Similarity	99.6%	Pred. No. 1.3e-120;		
Matches 227;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	ILGGREAEHARPYMASVOLNGAHLGCGVLVAEQWVLSAAHCLDEADAGKVQVLLGAHSL	60	
Db	11	ILGGREAEHARPYMASVOLNGAHLGCGVLVAEQWVLSAAHCLDEADAGKVQVLLGAHSL	70	
QY	61	SOPEPSKRLYDVLRAVPHPSQPDTHDHLQLSEKATLGPVAVRPLPWORVDRDVPAG	120	
Db	71	SOPEPSKRLYDVLRAVPHPSQPDTHDHLQLSEKATLGPVAVRPLPWORVDRDVPAG	130	
QY	121	TLCDVAGWGTNHNAGRRPDSLQHVLLPVLDRATCNRRTHDGAITERLMCAESNRRDSCK	180	
Db	131	TLCDVAGWGTNHNAGRRPDSLQHVLLPVLDRATCNRRTHDGAITERLMCAESNRRDSCK	190	
QY	181	GDSSGGLVCGGVLEGGVVTSGSRVCGNRKPKGIYTRVASYAAWIDSVLA	228	
Db	191	GDSSGGLVCGGVLEGGVVTSGSRVCGNRKPKGIYTRVASYAAWIDSVLA	238	
RESULT 2	5223425-4			
; Patent No. 5223425				
; APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,				
; BARRY M.; WHITE, TYLER R.				
; TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT				
; D ACTIVITY				
; NUMBER OF SEQUENCES: 19				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/07/277,963				
; FILING DATE: 30-NOV-1988				

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 34,203
; FILING DATE: 02-APR-1987
; SEQ ID NO:4:
; LENGTH: 250
5223425-4

Query Match      99.5%; Score 1205; DB 6; Length 250;
Best Local Similarity 99.6%; Pred. No. 1.4e-120;
Matches 227; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILGGREAEAHARPYMASVOLNGAHLGGVLAQWVLSAAHCLDEADGKQVLLGAHSL 60
Db 23 ILGGREAEAHARPYMASVOLNGAHLGGVLAQWVLSAAHCLDEADGKQVLLGAHSL 82
QY 61 SQPEPSKRLDYDLRAVPHDPDSQDPTIDHLLQLSEKATLGPVAPRPLPWQRVDRDVPAG 120
Db 83 SQPEPSKRLDYDLRAVPHDPDSQDPTIDHLLQLSEKATLGPVAPRPLPWQRVDRDVPAG 142
QY 121 TLCDVAGWGINVHAGRRPDSLOHVLPLVLDRTACNRTHHDGAIATERLMAESNRDSC 180
D. 143 TLCDVAGWGINVHAGRRPDSLOHVLPLVLDRTACNRTHHDGAIATERLMAESNRDSC 202
QY 181 GDSGGLVCGGVLEGVVTSGSRVCGNRKPKGIYTRVASYAAWIDSVLA 228
Db 203 GDSGGLVCGGVLEGVVTSGSRVCGNRKPKGIYTRVASYAAWIDSVLA 250

RESULT 3
5223425-8
; Patent No. 5223425
; APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,
; BARRY M.; WHITE, TYLER R.
; TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT
; D ACTIVITY
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/277,963
; FILING DATE: 30-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 34,203
; FILING DATE: 02-APR-1987
; SEQ ID NO:8:
; LENGTH: 253
5223425-8

Query Match      99.2%; Score 1201; DB 6; Length 253;
Best Local Similarity 99.1%; Pred. No. 3.9e-120;
Matches 226; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILGGREAEAHARPYMASVOLNGAHLGGVLAQWVLSAAHCLDEADGKQVLLGAHSL 60
Db 26 ILGGREAEAHARPYMASVOLNGAHLGGVLAQWVLSAAHCLDEADGKQVLLGAHSL 85
QY 61 SQPEPSKRLDYDLRAVPHDPDSQDPTIDHLLQLSEKATLGPVAPRPLPWQRVDRDVPAG 120
Db 86 SQPEPSKRLDYDLRAVPHDPDSQDPTIDHLLQLSEKATLGPVAPRPLPWQRVDRDVPAG 145
QY 121 TLCDVAGWGINVHAGRRPDSLOHVLPLVLDRTACNRTHHDGAIATERLMAESNRDSC 180
Db 146 TLCDVAGWGINVHAGRRPDSLOHVLPLVLDRTACNRTHHDGAIATERLMAESNRDSC 205
QY 181 GDSGGLVCGGVLEGVVTSGSRVCGNRKPKGIYTRVASYAAWIDSVLA 228
Db 206 GDSGGLVCGGVLEGVVTSGSRVCGNRKPKGIYTRVASYAAWIDSVLA 253

RESULT 4
US-08-944-483-44
; Sequence 44, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
```

```
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
US-08-944-483-44

Query Match      98.9%; Score 1198; DB 4; Length 228;
Best Local Similarity 98.7%; Pred. No. 7e-120;
Matches 225; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILGGREAEAHARPYMASVOLNGAHLGGVLAQWVLSAAHCLDEADGKQVLLGAHSL 60
Db 1 MLGGREAEAHARPYMASVOLNGAHLGGVLAQWVLSAAHCLDEADGKQVLLGAHSL 60
QY 61 SQPEPSKRLDYDLRAVPHDPDSQDPTIDHLLQLSEKATLGPVAPRPLPWQRVDRDVPAG 120
Db 61 SQPEPSKRLDYDLRAVPHDPDSQDPTIDHLLQLSEKATLGPVAPRPLPWQRVDRDVPAG 120
QY 121 TLCDVAGWGINVHAGRRPDSLOHVLPLVLDRTACNRTHHDGAIATERLMAESNRDSC 180
Db 121 TLCDVAGWGINVHAGRRPDSLOHVLPLVLDRTACNRTHHDGAIATERLMAESNRDSC 180
QY 181 GDSGGLVCGGVLEGVVTSGSRVCGNRKPKGIYTRVASYAAWIDSVLA 228
Db 181 GDSGGLVCGGVLEGVVTSGSRVCGNRKPKGIYTRVASYAAWIDSVLA 228

RESULT 5
5223425-6
; Patent No. 5223425
; APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,
; BARRY M.; WHITE, TYLER R.
```

; TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT

; D ACTIVITY  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/277,963  
; FILING DATE: 30-NOV-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 34,203  
; FILING DATE: 02-APR-1987  
; SEQ ID NO:6:  
; LENGTH: 222  
5223425-6

Query Match 86.8%; Score 1051.5; DB 6; Length 222;  
Best Local Similarity 89.9%; Pred. No. 2.8e-104;  
Matches 205; Conservative 4; Mismatches 12; Indels 7; Gaps 3;

Qy 1 ILGREAHAHAPYMASVOLNCAHLCGGVLAQWVLSAAHCLDEADAGKQVLLGAHSL 60  
1 ILGREAHAHAPYMASVOLNCAHLCGGVLAQWVLSAAHCLDEADAGKQVLLGAHSL 60  
61 SQPEPSKRLDYVLRVPHPSQPTDIDHLLQLSEKATLGPVRLPWQVDRDVP 119  
Db 61 PQPEPXXITIEVLRVPHPSQPTDIDHLLQLSEKATLGPVRLPWQVDRDVP 120  
Qy 120 GTLCDVAGWGVVTHAGRRPDSLOHVLPLVDRATCNRRTHHDGATTELMAESNRDSC 179  
Db 121 GTLCDVAGWGVVTHAGRRPDSLOHVLPLVDRATCNRRTHHDGATTELMAESNRDSC 174  
Qy 180 KDGSGPLVCGVLEGGVTSRVCGRNKKPGIYTRVASYAAWIDSVL 227  
Db 175 KDGSGPLVCGVLEGGVTSRVCGRNKKPGIYTRVASYAAWIDSVL 222

RESULT 6  
5223425-2

; Patent No. 5223425  
; APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,  
; BARRY M.; WHITE, TYLER R.  
; TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT

; D ACTIVITY  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/277,963  
; FILING DATE: 30-NOV-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 34,203  
; FILING DATE: 02-APR-1987  
; ID NO:2:  
; LENGTH: 259  
5223425-2

Query Match 69.7%; Score 844.5; DB 6; Length 259;  
Best Local Similarity 66.5%; Pred. No. 4e-82;  
Matches 151; Conservative 33; Mismatches 42; Indels 1; Gaps 1;

Qy 1 ILGREAHAHAPYMASVOLNCAHLCGGVLAQWVLSAAHCLDEADAGKQVLLGAHSL 59  
Db 26 ILGREAHAHAPYMASVOLNCAHLCGGVLAQWVLSAAHCLDEADAGKQVLLGAHSL 85  
Qy 60 LSQPEPSKRLDYVLRVPHPSQPTDIDHLLQLSEKATLGPVRLPWQVDRDVP 119  
Db 86 LSQPEPSKRLDYVLRVPHPSQPTDIDHLLQLSEKATLGPVRLPWQVDRDVP 145  
Qy 120 GTLCDVAGWGVVTHAGRRPDSLOHVLPLVDRATCNRRTHHDGATTELMAESNRDSC 179  
Db 146 GTLCDVAGWGVVTHAGRRPDSLOHVLPLVDRATCNRRTHHDGATTELMAESNRDSC 205  
Qy 180 KDGSGPLVCGVLEGGVTSRVCGRNKKPGIYTRVASYAAWIDSV 226  
Db 206 KDGSGPLVCGVLEGGVTSRVCGRNKKPGIYTRVASYAAWIDSV 252

RESULT 7  
5223425-10

; Patent No. 5223425  
; APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,  
; BARRY M.; WHITE, TYLER R.  
; TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT

; D ACTIVITY  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/277,963  
; FILING DATE: 30-NOV-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 34,203  
; FILING DATE: 02-APR-1987  
; SEQ ID NO:10:  
; LENGTH: 260  
5223425-10

Query Match 69.7%; Score 844; DB 6; Length 260;  
Best Local Similarity 66.2%; Pred. No. 4.5e-82;  
Matches 151; Conservative 33; Mismatches 42; Indels 2; Gaps 1;

Qy 1 ILGREAHAHAPYMASVOLNCAHLCGGVLAQWVLSAAHCLDEADAGKQVLLGAHSL 58  
Db 26 ILGREAHAHAPYMASVOLNCAHLCGGVLAQWVLSAAHCLDEADAGKQVLLGAHSL 85  
Qy 59 LSQPEPSKRLDYVLRVPHPSQPTDIDHLLQLSEKATLGPVRLPWQVDRDVP 118  
Db 86 LSQPEPSKRLDYVLRVPHPSQPTDIDHLLQLSEKATLGPVRLPWQVDRDVP 145  
Qy 119 GTLCDVAGWGVVTHAGRRPDSLOHVLPLVDRATCNRRTHHDGATTELMAESNRDSC 178  
Db 146 GTLCDVAGWGVVTHAGRRPDSLOHVLPLVDRATCNRRTHHDGATTELMAESNRDSC 205  
Qy 179 KDGSGPLVCGVLEGGVTSRVCGRNKKPGIYTRVASYAAWIDSV 226  
Db 206 KDGSGPLVCGVLEGGVTSRVCGRNKKPGIYTRVASYAAWIDSV 253

RESULT 8  
US-08-944-483-45

; Sequence 45, Application US/08944483  
; Patent No. 6232456  
; GENERAL INFORMATION:

; APPLICANT: COHEN, MAURICE  
; APPLICANT: COLPITTS, TRACEY L.  
; APPLICANT: FRIEDMAN, PAULA N.  
; APPLICANT: GRANADOS, EDWARD N.  
; APPLICANT: KLASS, MICHAEL R.  
; APPLICANT: RUSSELL, JOHN C.  
; APPLICANT: STEWART, KENT D.  
; APPLICANT: STROUPE, STEVEN D.  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS  
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
; TITLE OF INVENTION: OF THE PROSTATE  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/944,483  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
US-08-944-483-45

Query Match 38.98; Score 471.5; DB 4; Length 232;
Best Local Similarity 43.38; Pred. No. 2e-42;
Matches 100; Conservative 41; Mismatches 81; Indels 9; Gaps 5;

QY 1 ILGREGAEAHARPYMASVOLNGAHLGCGVLVAEQWVLSAAHCLDEADAADGKVQVLLGAHSL 60
Db 1 ILGREGAEAHARPYMASVOLNGAHLGCGVLVAEQWVLSAAHCLDEADAADGKVQVLLGAHSL 59
QY 61 SQPEPSKRLVDVLAHVPDPSQ--DTIDHLLQLLSEKATLGPVAVRPLPQWRVDRDVP 119
Db 60 DSPGLT---PHIKAAIQHPKYPKVPALNDLALLQDQKVPKPSRTIRPLALPSKQVVA 116
QY 120 GTLCDVAGWGVNIHAGRPDSLOHVLPLVDLRATCNRTHDGAITELCAESNRDS- 178
Db 117 GTRCSMAGWGLTHGGRLSRVRLDQLVLDTRCMNNSRWNGSLSPMVCVLAADSKDQA 176
QY 179 -CKDGGSGPLVC--GGVLEGVVTSGSRVCGNRKKPGIVTRVASYAAWIDSV 226
Db 177 PKGDSGGLVCGGRVLAGVLSSFRVCTDIKPPVATAVAPVSWIRKV 227

RESULT 9
US-08-944-483-46
; Sequence 46, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLETTIS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; TITLE OF INVENTION: OF THE PROSTATE
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
US-08-944-483-46

Query Match 37.78; Score 456.5; DB 4; Length 234;
Best Local Similarity 41.58; Pred. No. 7.9e-41;
Matches 95; Conservative 44; Mismatches 81; Indels 9; Gaps 5;

QY 1 ILGREGAEAHARPYMASVOLNGAHLGCGVLVAEQWVLSAAHCLDEADAADGKVQVLLGAHSL 60
Db 1 IIGNEVTPHSRPRYVLLSLDRKTCAGALIAKDWLTAHC-----NLNKRQVLLGAHSI 57
QY 61 SQPEPSKRLVDVLAHVPDPSQPTIDHLLQLLSEKATLGPVAVRPLPQWRVDRDVP 120
Db 58 TREPTQIMLVKKEFPYPCYDPATREGDLKLQLETEKAKINKYVTLHLPKGDDVKPG 117
QY 121 TLCDVAGWGVNIHAGRPDSLOHVLPLVDLRATCNRTHH--DGAITERLMCAESNR--R 176
Db 118 TMCQVAGWGRTHNSASWSDTLREVNITIIDRKVCNDRNHNFNFNIGNMVYVIGNMGV 177
QY 177 DSCGDSGGLVCGGVLEGVVTSG-SRVCGNRKKPGIVTRVA-SYAAMI 223
Db 178 DSCNGDSGSPLLCEGVFRGVTSGFLENKCGDPRGPGVYLLSKKHLAWI 226

RESULT 10
US-07-990-301A-4
; Sequence 4, Application US/07990301A
; Patent No. 5500344
; GENERAL INFORMATION:
; APPLICANT: SAYERS, THOMAS; WILTROUT, THERESA A.;
; APPLICANT: POWERS, JAMES C.; SMYTH, MARK J.; SOWDER, RAYMOND;
; APPLICANT: HENDERSON, LOUIS E.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVE.
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT # 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/990,301A
; FILING DATE: 03-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2025-4055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
```

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-758-4800
: TELEFAX: 212-751-6849
: TELEX: 421792
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 258
: TYPE: AMINO ACID
: TOPOLOGY: UNKNOWN
: MOLECULE TYPE:
: DESCRIPTION: PROTEIN
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: RAT
: STRAIN:
: INDIVIDUAL ISOLATE:
: DEVELOPMENTAL STAGE:
: HARLOTYPE:
: TISSUE TYPE:
: CELL TYPE: CYTOLYTIC GRANULE
: CELL LINE: RNK-16 LGL
: ORGANELLE:
:
US-07-990-301A-2

Query Match          36.7%; Score 445; DB 1; Length 258;
Best Local Similarity 39.8%; Pred. No. 1.5e-39;
Matches 92; Conservative 48; Mismatches 85; Indels 6; Gaps

QY   1 ILGGRAEAHARYMASVOLNGAHLCGGVLYAEOWVLSSAAHCLEDADCKVQVLLCAHSL 60
      | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : |||
Db    21 IIGGREAVPSRPMYSVLQNTKSHMCGGVLVHQKWYLTAACHLSEPLQ-QLKLVFLGHL 79
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   61 SQPEPKRLDYLRVAPHPDSQPDTHDHLILLQLSEKATLGPAVRPLPWQRVDRD-VAP 119
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    80 HDQDFCLTFYIKQAHKPCYNL-KYENDALLUKLDGRVKPKNKWKPALPRKPDKPAE 138
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   120 GTLCDVAGWGIYNHAGRRPDSLQHVLPLVDLRATCNRRTHHDGAITERLMCAESNRDRS- 178
      | : |||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : |||
Db    139 GSRCSTAGWGITHQGQLAKLSQELDLRLDTRMCMNSRFNMGVLTDSMLCLKAGAKGOA 198
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   179 -CKGDGGPLVCC-GVLEGVVTSRGVCGRKKPGIYTRVASVAWIDSVL 227
      | ||||||||| | : ||| | : ||| | : ||| | : ||| | : ||| | : |||
Db    199 PKRGDGGPLVCCGKVGDGILSFSSKNCTDIKPTVATAVPISSWIRKVI -249
      | ||||||||| | : ||| | : ||| | : ||| | : ||| | : ||| | : |||

RESULT 12
US-09-008-271A-7
: Sequence 7, Application US/09008271A
: Patent No. 6203979
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: Hillman, Jennifer L.
: Yue, Henry
: Guegler, Karl J.
: Corley, Neil C.
: Tang, Tom Y.
: Shah, Purvi
: TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Dr.
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/008.271A

```



TELEX: 846169  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 260 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-070-526-2

Search completed: May 1, 2003, 22:09:27  
Job time : 30 secs

Query Match 35.8%; Score 433.5; DB 3; Length 260;  
Best Local Similarity 39.5%; Pred. No. 2.6e-38;  
Matches 90; Conservative 42; Mismatches 83; Indels 13; Gaps 6;

Qy 1 ILGGREAHAHPYMASVOLNGAHLCCGVLVAEQWLSAAHCLDEADAGKQVVLGAHSL 60  
Db 33 VLGGHECPHSPQWQAALFQGOQLCCGVLVGGWNLTAACHKKP----KYTVRLGDHSL 88  
Qy 61 SQPEPSKRLVDLRAVPHP---DSQPTDIDHDLILLQLSEKATLGPVAPRPLPWQRVDRD 117  
89 QNKDGPQEIIPVVGSIHPHCYNSSDVEDHNDLMLQLDRQASLGSKVKPI--SLADHCT 146  
Qy 118 APTGLCDVAGMGIVNHAGRR-PSLQHLVLLPVLDRATCNRRTTHDGAITERLMCAESNR- 175  
Db 147 QPGQKCTVSGWGTVTSPRENFPTLNCAEVKIFPKKC--EDAYPGQITDGMVCAGSKG 204  
Qy 176 RDSCKGDSGGPLVCGVLEGVVTSGRVCGRNKKPGIYTRVASTAAWI 223  
Db 205 ADTCQDGGSGPLVCDGALQGITSMGSDPCGRSDKPGVYTNICRYLDWI 252

RESULT 15

US-09-386-642-13  
Sequence 13, Application US/09386642  
Patent No. 6420157  
GENERAL INFORMATION:  
APPLICANT: Darrow, Andrew  
APPLICANT: Qi, Jensen  
APPLICANT: Andrade-Gordon, Patricia  
TITLE OF INVENTION: Zymogen Activation System  
FILE REFERENCE: ORT-1028  
CURRENT APPLICATION NUMBER: US/09/386.642  
CURRENT FILING DATE: 1999-08-31  
NUMBER OF SEQ ID NOS: 60  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 13  
LENGTH: 288  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Fusion gene  
OTHER INFORMATION: with homo sapien serine protease catalytic domain  
US-09-386-642-13

Query Match 34.8%; Score 421; DB 4; Length 288;  
Best Local Similarity 38.6%; Pred. No. 6.4e-37;  
Matches 90; Conservative 43; Mismatches 96; Indels 14; Gaps 7;

Qy 1 ILGGREAHAHPYMASVOLNGAHLCCGVLVAEQWLSAAHCLDEADAGKQVVLGAHSL 59  
Db 52 IVGGYNCLEPHSPQWQAALFQGOQLCCGVLVGGWNLTAACHKKP----KYTVRLGDHSL 107  
Qy 60 LSQPEPSKRLVDLRAVPHP---DSQPTDIDHDLILLQLSEKATLGPVAPRPLPWQRVDRD 116  
Db 108 LONKDGPEQEIIPVVGSIHPHCYNSSDVEDHNDLMLQLDRQASLGSKVKPI--SLADHC 165  
Qy 117 VAPGTLCDVAGMGIVNHAGRR-PSLQHLVLLPVLDRATCNRRTTHDGAITERLMCAESNR 175  
Db 166 TQPGQKCTVSGWGTVTSPRENFPTLNCAEVKIFPKKC--EDAYPGQITDGMVCAGSKG 223  
Qy 176 -RDSCKGDSGGPLVCGVLEGVVTSGRVCGRNKKPGIYTRVASTAAWIDSVL 227  
Db 224 GADTCQDGGSGPLVCDGALQGITSMGSDPCGRSDKPGVYTNICRYLDWIKKII 276

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 22:01:05 ; Search time 76 Seconds  
(without alignments)  
399.752 Million cell updates/sec

Title: US-09-821-255-2  
Perfect score: 1211  
Sequence: 1 ILGGRAEAHARPMASVQL.....KPGIYTRVASYAAWIDSVLA 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

. 1 number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1202	99.3	250	11	AA05772
2	1201	99.2	250	11	AA05421
3	1098.5	90.7	212	20	AA28590
4	1056	87.2	233	20	AA28591
5	456.5	37.7	262	8	AA071672
6	456.5	37.7	264	18	AA027296
7	451	37.2	284	23	AA082754
8	447.5	37.0	232	15	AA057578
9	447	36.9	283	20	AA28640
10	447	36.9	283	21	AA099371

11	447	36.9	283	21	AA070026
12	447	36.9	283	22	AA029159
13	447	36.9	283	22	AA02497
14	447	36.9	283	22	AA06120
15	446	36.8	283	22	AA070778
16	445	36.7	258	15	AA055757
17	440.5	36.4	260	20	AA06438
18	437.5	36.1	260	20	AA01744
19	437.5	36.1	260	20	AA032852
20	437.5	36.1	260	20	AA03220
21	437.5	36.1	260	21	AA01322
22	437.5	36.1	260	21	AA044300
23	437.5	36.1	260	21	AA011131
24	437.5	36.1	260	22	AB023373
25	437.5	36.1	260	22	AA012369
26	437.5	36.1	260	22	AA053087
27	437.5	36.1	260	23	AB05458
28	437.5	36.1	260	23	AB04852
29	437.5	36.1	260	23	AA081959
30	437.5	36.1	275	21	AA021311
31	437.5	36.1	305	20	AA032853
32	437.5	36.1	305	21	AA037985
33	437.5	36.1	315	23	AB041332
34	433.5	35.8	244	15	AA04532
35	433.5	35.8	244	18	AA022985
36	433.5	35.8	244	19	AA051006
37	433.5	35.8	244	21	AA021323
38	433.5	35.8	260	20	AA087703
39	432.5	35.7	240	23	AB04644
40	431.5	35.6	239	21	AA021310
41	424	35.0	249	23	AA021441
42	421.5	34.8	253	20	AA080025
43	421	34.8	288	21	AA036482
44	421	34.8	288	22	AA067542
45	419.5	34.6	260	17	AA010694

ALIGNMENTS

RESULT 1	
AA05772	
ID	AA05772 standard; protein; 250 AA.
AC	
XX	AA05772;
XX	
DT	05-NOV-1990 (first entry)
XX	
DE	Human adipsin gene product from the clone phg31.
XX	
KW	Adipsin; complement D; obesity; lupus erythematosis;
KW	rheumatoid arthritis; ds.
XX	
PN	W09006365-A.
XX	
PD	14-JUN-1990.
XX	
PF	
XX	21-NOV-1989; 89WO-0005374.
XX	
PR	
XX	30-NOV-1988; 88US-0277963.
PA	(BETH-) BETH ISRAEL HOSPITA.
PA	(DANA-) DANA-FARBER CANCER INST.
XX	(META-) METABOLIC BIOSYSTEMS INC.
PI	Flier JS, Spiegelman BM, Rosen BM, White RT;
XX	
DR	WPI; 1990-209777/27.
DR	N-PSDB; AAQ05149.
XX	
PT	Human protein having adipsin and complement D activity -
PT	used in diagnosis and control of metabolically caused obesity
PT	and for treating infection.



```
XX
PS Disclosure; : p: English.
CC
CC Abs raised to the gene product may be used in diagnosis of
CC metabolically caused obesity. The protein may also be used to
CC treat and prevent obesity and bacterial, viral, parasitic and
CC neoplastic cell infection.
XX
SQ Sequence 250 AA;

Query Match 99.3%; Score 1202; DB 11; Length 250;
Best Local Similarity 99.1%; Pred. No. 1.4e-105;
Matches 226; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILGREAHAHARPYMASVOLNGAHLGGVILVAEQWVLSAAHCLDEADAADGKVQLLGAHSL 60
DB 23 ILGREAHAHARPYMASVOLNGAHLGGVILVAEQWVLSAAHCLDEADAADGKVQLLGAHSL 82
QY 61 SQPEPSKRLYDVLRAVPHDPDSQPDTHDHLQLLQSEKATLGPVAVRPLPWQVRDVRVAPG 120
DB 83 SQPEPSKRLYDVLRAVPHDPDSQPDTHDHLQLLQSEKATLGPVAVRPLPWQVRDVRVAPG 142
QY 121 TLCDVAGWGIVNAGRRPDSLSQHVLLPVLDRTATNRRTHDGAITERLMCAESNRRDSC 180
DB 143 TLCDVAGWGIVNAGRRPDSLSQHVLLPVLDRTATNRRTHDGAITERLMCAESNRRDSC 202
QY 181 GDSGGPLVCGGVLEGGVYVTSGRVCGNKKKPGIYTRVASYAAWIDSVLA 228
DB 203 GDSGGPLVCGGVLEGGVYVTSGRVCGNKKKPGIYTRVASYAAWIDSVLA 250

RESULT 2
AAR05421
ID AAR05421 standard; protein; 250 AA.
AC AAR05421;
XX
XX 30-JUL-1990 (first entry)
XX
XX Human adipsin/D encoded by a cDNA.
XX
XX CAT; hybrid protein; Human adipsin/D.
XX
XX Homo sapiens.
XX
XX WO9001540-A.
XX
XX 22-FEB-1990.
XX
XX 19-AUG-1989; 89WO-US03417.
XX
XX 11-AUG-1988; 88US-0231224.
XX
XX (CALB-) CALIF BIOTECHN INC.
XX
XX Hilliker S.; White R;
XX
XX WPI; 1990-083499/11.
XX
XX N-PSDB; AAQ03566.
XX
XX heterologous protein expression on prokaryotic host -
XX using 3' truncated chloramphenicol acetyl transferase gene to
XX stably express hybrid protein.
XX
XX Example; Fig 10; 67pp; English.
XX
XX When inserted into an expression vector, pTIPCAT 72,
XX the construct gave 10-15% levels of fusion protein upon
XX induction in W3110 cells.
XX
XX See also AAQ03557 to AAQ03566; and AAQ04767.
XX
XX Sequence 250 AA;

Query Match 99.1%; Score 1201; DB 11; Length 250;
Best Local Similarity 99.1%; Pred. No. 1.8e-105;
Matches 226; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILGREAHAHARPYMASVOLNGAHLGGVILVAEQWVLSAAHCLDEADAADGKVQLLGAHSL 60
DB 23 ILGREAHAHARPYMASVOLNGAHLGGVILVAEQWVLSAAHCLDEADAADGKVQLLGAHSL 82
QY 61 SQPEPSKRLYDVLRAVPHDPDSQPDTHDHLQLLQSEKATLGPVAVRPLPWQVRDVRVAPG 120
DB 83 SQPEPSKRLYDVLRAVPHDPDSQPDTHDHLQLLQSEKATLGPVAVRPLPWQVRDVRVAPG 142
QY 121 TLCDVAGWGIVNAGRRPDSLSQHVLLPVLDRTATNRRTHDGAITERLMCAESNRRDSC 180
DB 143 TLCDVAGWGIVNAGRRPDSLSQHVLLPVLDRTATNRRTHDGAITERLMCAESNRRDSC 202
QY 181 GDSGGPLVCGGVLEGGVYVTSGRVCGNKKKPGIYTRVASYAAWIDSVLA 228
DB 203 GDSGGPLVCGGVLEGGVYVTSGRVCGNKKKPGIYTRVASYAAWIDSVLA 250

RESULT 3
AAY28590
ID AAY28590 standard; Protein; 212 AA.
XX
XX AAY28590;
XX
XX 26-NOV-1999 (first entry)
XX
XX Human Factor D.
XX
XX Factor D; alternative complement pathway; serine protease; asthma;
XX cardiopulmonary bypass; tissue damage; Alzheimer's disease;
XX multiple sclerosis; acute myocardial infarction; Crohn's disease;
XX transplant rejection; systemic lupus erythematosus; inflammation.
XX
XX Homo sapiens.
XX
XX WO9942133-A1.
XX
XX 26-AUG-1999.
XX
XX 19-FEB-1999; 99WO-US03566.
XX
XX 20-FEB-1998; 98US-0075328.
XX
XX (TANO-) TANOX INC.
XX
XX Fung MSC, Sun BNC, Sun CRY;
XX
XX WPI; 1999-527424/44.
XX
XX N-PSDB; AA206673.
XX
XX Factor D inhibitors useful for treating pathological inflammation
XX and autoimmune diseases
XX
XX Example 1; Page 93-94; 99pp; English.
XX
XX This sequence is human factor D. Factor D is a highly specific serine
XX protease essential for activation of the alternative complement pathway.
XX The plasma concentration of factor D in humans is very low making it a
XX suitable target for inhibition. It is the limiting enzyme for the
XX activation of the alternative complement pathway. Factor D is used to
XX generate inhibitors of complement activation which bind factor D at a
XX molar ratio of at least 1.5:1 and less than 80:1 (inhibitor:factor D).
XX Factor D was used to design a monoclonal antibody MAB 166-32 which binds
XX to factor D and blocks its ability to activate complement. The
XX anti-factor D molecules can function to inhibit in vivo complement
XX activation and/or the alternative complement pathway and inflammatory
XX events which occur. These include the recruitment and activation of
XX macrophages, neutrophils, platelets, and mast cells, oedema, and tissue
XX damage. The inhibitors of complement activation may be used to treat
XX diseases or conditions that are mediated by excessive or uncontrolled
```

CC activation of the complement system. These inhibitors are also  
 CC administered for treating complement-mediated conditions associated with  
 CC cardiopulmonary bypass. Diseases or conditions that are mediated by  
 CC excessive or uncontrolled activation of the complement systems include  
 CC tissue damage due to ischaemia-reperfusion following acute myocardial  
 CC infarction, aneurysm, stroke, haemorrhagic shock, crush injury, multiple  
 CC organ failure; inflammatory disorders, e.g. burns, endotoxaemia and  
 CC septic shock; haemodialysis, severe asthma, Crohn's disease; transplant  
 CC rejection; and adverse drug reaction. Autoimmune disorders can also be  
 CC treated such as systemic lupus erythematosus, Alzheimer's disease and  
 CC multiple sclerosis. The anti-factor D molecules can also be used  
 CC diagnostically to determine the presence of or quantity of factor D in a  
 CC tissue specimen or a body fluid sample.

SQ Sequence 212 AA;  
 Query Match 90.7%; Score 1098.5; DB 20; Length 212;  
 Best Local Similarity 92.5%; Pred. No. 7.2e-96;  
 Matches 210; Conservative 1; Mismatches 1; Indels 15; Gaps 1;  
 1 ILGGREAEHARPYMASVOLNGAHLGCGVLVAEQWVLSAAHCLDAADGKVQVLLGAHSL 60  
 1 ILGGREAEHARPYMASVOLNGAHLGCGVLVAEQWVLSAAHCLDAADGKVQVLLGAHSL 60  
 61 SOPEPSKRLVDLRAVPHDPDSQPTIDHDLILLLOLSEKATLGPVRLPWQVRDVA PG 120  
 61 SOPEPSKRLVDLRAVPHDPDSQPTIDHDLILLLOLSEKATLGPVRLPWQVRDVA PG 120  
 121 TLCDVAGWGIYNHAGRRPDSLQHVLLPVLDRATCNRRTHDGAITERLMCAESNRDSC 180  
 121 TLCDVAGWGIYNHAGRRPDSLQHVLLPVLDRATCNRRTHDGAITERLMCAESNRDSC 180  
 181 GDSGGLVCGGVLEGVVTSGSRVCGNRKPKGIYTRVASYAAWIDSVL 227  
 166 GDSGGLVCGGVLEGVVTSGSRVCGNRKPKGIYTRVASYAAWIDSVL 212

## RESULT 4

AAY28591  
 ID AAY28591 standard; Protein: 233 AA.

AC AAY28591;

26-NOV-1999 (first entry)

Pig Factor D.

Factor D; alternative complement pathway; serine protease; asthma;  
 cardiopulmonary bypass; tissue damage; Alzheimer's disease;  
 multiple sclerosis; acute myocardial infarction; Crohn's disease;  
 transplant rejection; systemic lupus erythematosus; inflammation.

Sus scrofa.

WO9942133-A1.

26-AUG-1999.

19-FEB-1999; 99WO-US03566.

20-FEB-1998; 98US-0075328.

(TANO-) TANOX INC.

Fung MSC, Sun BNC, Sun CRY;

WPI; 1999-527424/44.

N-PSDB; AA206674.

Factor D inhibitors useful for treating pathological inflammation  
 and autoimmune diseases

Example 8; Page 95; 99pp; English.

xx This sequence is the amino acid sequence of pig factor D. Factor D is a  
 CC highly specific serine protease essential for activation of the  
 CC alternative complement pathway. The plasma concentration of factor D in  
 CC humans is very low, making it a suitable target for inhibition. It is the  
 CC limiting enzyme for the activation of the alternative complement pathway.  
 CC Factor D is used to generate inhibitors of complement activation which  
 CC bind factor D at a molar ratio of at least 1.5:1 and less than 80:1  
 CC (inhibitor:factor D). Human factor D was used to design a monoclonal  
 CC antibody MAB 166-32 which binds to factor D and blocks its ability to  
 CC activate complement. Pig factor D is used in the construction of human  
 CC factor D mutants to be used for epitope mapping of MAB 166-32. The  
 CC anti-factor D molecules can function to inhibit in vivo complement  
 CC activation and/or the alternative complement pathway and inflammatory  
 CC events which occur. These include the recruitment and activation of  
 CC macrophages, neutrophils, platelets, and mast cells, oedema, and tissue  
 CC damage. The inhibitors of complement activation may be used to treat  
 CC diseases or conditions that are mediated by excessive or uncontrolled  
 CC activation of the complement system. These inhibitors are also  
 CC administered for treating complement-mediated conditions associated with  
 CC cardiopulmonary bypass. Diseases or conditions that are mediated by  
 CC excessive or uncontrolled activation of the complement systems include  
 CC tissue damage due to ischaemia-reperfusion following acute myocardial  
 CC infarction, aneurysm, stroke, haemorrhagic shock, crush injury, multiple  
 CC organ failure; inflammatory disorders, e.g. burns, endotoxaemia and  
 CC septic shock; haemodialysis, severe asthma, Crohn's disease; transplant  
 CC rejection; and adverse drug reaction. Autoimmune disorders can also be  
 CC treated such as systemic lupus erythematosus, Alzheimer's disease and  
 CC multiple sclerosis. The anti-factor D molecules can also be used  
 CC diagnostically to determine the presence of or quantity of factor D in a  
 CC tissue specimen or a body fluid sample.

SQ Sequence 233 AA;

Query Match 87.2%; Score 1056; DB 20; Length 233;

Best Local Similarity 84.2%; Pred. No. 8.5e-92;

Matches 192; Conservative 18; Mismatches 18; Indels 0; Gaps 0;

QY 1 ILGGREAEHARPYMASVOLNGAHLGCGVLVAEQWVLSAAHCLDAADGKVQVLLGAHSL 60  
 DB 1 ILGGREAEHARPYMASVOLNGAHLGCGVLVAEQWVLSAAHCLDAADGKVQVLLGAHSL 60

QY 61 SOPEPSKRLVDLRAVPHDPDSQPTIDHDLILLLOLSEKATLGPVRLPWQVRDVA PG 120  
 DB 61 SOPEPSKRLVDLRAVPHDPDSQPTIDHDLILLLOLSEKATLGPVRLPWQVRDVA PG 120

QY 121 TLCDVAGWGIYNHAGRRPDSLQHVLLPVLDRATCNRRTHDGAITERLMCAESNRDSC 180  
 DB 121 TLCDVAGWGIYNHAGRRPDSLQHVLLPVLDRATCNRRTHDGAITERLMCAESNRDSC 180

QY 181 GDSGGLVCGGVLEGVVTSGSRVCGNRKPKGIYTRVASYAAWIDSVL 228  
 DB 181 GDSGGLVCGGVLEGVVTSGSRVCGNRKPKGIYTRVASYAAWIDSVL 228

## RESULT 5

AAP71672

ID AAP71672 standard; Protein: 262 AA.

AC AAP71672;

10-JUN-1991 (first entry)

Human serine protease.

Serine protease; assay; antibodies; immunisation; HSP; diagnosis.

Homo sapiens.

Key Location/Qualifiers  
 Peptide 1..28  
 /label= sig\_peptide  
 Protein 28

/label= mat\_protein

FT XX  
PN XX  
XX EP245051-A.  
XX 11-NOV-1987.  
XX 01-MAY-1987; 87EP-0303945.  
XX 31-DEC-1986; 86US-0948248.  
XX 06-MAY-1986; 86US-0860085.  
XX 08-MAY-1986; 86US-0861221.  
XX (STRD ) LELAND STANFORD JR UNIV.  
XX Weissman IL, Gershenfeld HK;  
XX WPI; 1987-315213/45.  
XX N-PSDB; AAN71407.  
XX New pure human serine protease and fragments - used as labels in  
P assays and for prodn. of antibodies for passive immunisation  
against immune disorders.  
XX Disclosure; Fig 1; 7pp; English.  
XX Amino acid homology within the active enzyme to the mouse protease is  
71% with 77% at the DNA level. The overall homology is 72% when the  
complete coding region and the 3' UTR are included. The amino acids  
of the charge-relay system are His41, Asp86 and Ser184. The acidic  
residue Asp178 determines substrate specificity for Lys or Arg.  
XX Asn142 is an Asn-linked carbohydrate site.  
XX The HSP is produced by activated killer cells. The enzyme acts in  
conjunction with other components of a killer cell to provide  
cytolytic capability.  
XX SQ Sequence 262 AA;

Query Match 37.7%; Score 456.5; DB 8; Length 262;  
Best Local Similarity 41.5%; Pred. No. 4.7e-35;  
Matches 95; Conservative 44; Mismatches 81; Indels 9; Gaps 5;

QY 1 ILGGREAEAHARPYMASVOLNGAHLGGVLAQWVLSAAHCLDEADGKGVVLLGAHSL 60  
DB 29 IIGNEVTPHSPRYMVLSSDRKTCACALLAKDWLTAAC---NLNKRQVILGAHSI 85  
QY 61 SQPEPSKRLDYVLRVPHDPSPQPTDIDHLLQLSEKATLGPVRLPWRVDRDVPAG 120  
DB 86 TREPTKQIMLVKKEFPYPCYDPATREGDLKLLQTEKAKINKYVITILHLPKGDVDPKPG 145  
Q 121 TLCDVAGGIVNHAGRRPDSLQHVLLPVLDRATCNRRTHH--DGAITERLMCAESNR--R 176  
DB 146 TMCQVAGHGRTHNSASCDTLREYNITIDRKVCNDRNHYNFVIGNMVMYCAAGSLRGR 205  
QY 177 DSCKGDGGPLVCGVLEGVVTSG-SRVCGNKRKPGIYTRVA-SYAAMI 223  
DB 206 DSCNGDSGPLCEGVFRTVSFGLNKCGRPGVGVVILLSKRLHAWI 254

RESULT 6  
AAW27296  
ID AAW27296 standard; Protein; 264 AA.

XX AAW27296;

XX 19-MAY-1998 (first entry)

XX Human H83-22 secreted protein.

XX Secreted protein; anti-inflammatory; immune stimulant; suppressant;  
KW human; cytokine; autoimmune disease; regulator; activin; inhibin;  
XX H83-22.

XX Homo sapiens.

XX Key Location/Qualifiers  
FH Protein 1..264  
FT /label= H83.22  
FT /note= "secreted protein"

XX W09740151-A2.  
XX 30-OCT-1997.  
XX 14-APR-1997; 97WO-US06125.  
XX 10-JAN-1997; 97US-0781226.  
XX 19-APR-1996; 96US-0635311.  
XX (GEMY ) GENETICS INST INC.

XX Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA;  
PI Spaulding V;  
XX WPI; 1997-535835/49.  
XX N-PSDB; AAT91308.

XX Nucleic acids encoding novel secreted proteins - useful e.g. as an  
PT anti-inflammatory, immune stimulant or suppressor, etc.  
XX Claim 24; Page 65-66; 81pp; English.

XX This sequence represents a novel secreted protein, H83-22, which is  
CC isolated from a clone, H83-22 (ATCC 98028), derived from a human PBMC  
CC cDNA library. The nucleic acid and encoded H83-22 secreted protein can be  
CC used for research purposes (as markers for tissues, molecular weight  
CC markers for gels, primers, probes, etc.), for nutrition (as C, N or  
CC carbohydrate source), as a cytokine for cell proliferation and  
CC differentiation activity, as immune stimulants or suppressors e.g.  
CC for viral, bacterial or fungal infections, for autoimmune diseases such  
CC as multiple sclerosis or systemic lupus erythematosus, to regulate  
CC haematopoiesis, for tissue growth, as an activin or inhibin, or having  
CC chemotactic, chemokinetic, haemostatic and thrombolytic, receptor/ligand,  
CC anti-inflammatory or tumour inhibitory activities.

XX SQ Sequence 264 AA;

Query Match 37.7%; Score 456.5; DB 18; Length 264;  
Best Local Similarity 42.7%; Pred. No. 4.7e-35;  
Matches 100; Conservative 40; Mismatches 85; Indels 9; Gaps 7;

QY 1 ILGGREAEAHARPYMASVOLNGAHLGGVLAQWVLSAAHCLDEADGKGVVLLGAHSL 59  
DB 27 IIGKEVSPHSRPFPMASIQYGHVHCVGGVLDIPQWVLTAAHCQYRFTKGQSPVVLGAHS 86  
QY 60 LSQPEPSKRLDYVLRVPHDPSPQPTDIDHLLQLSEKATLGPVRLPWRVDRDVPAG 119  
DB 87 LSKNEASKOTLEIKRFIPFSRVTSQPSQNDIMLVKLTAAKLNKHVKML-HIRSKTSLRS 145  
QY 130 GTLCDVAGGIVNHAGRRP-DSLQHVLLPVLDRATCNRRTHH--AITERLMCA--ESN 174  
DB 146 GTKCKVTGATDPSLRSDTLREYTVVLSKLCNQSQSYNGDPFITKDWVCAGDAKG 205  
QY 175 RDSCKGDGGPLVCGVLEGVVTSGSRVCGNKRKPGIYTRVA-SYAAMIDSVL 227  
DB 206 QKDSCKGDGGPLCKGVFHAIV-SGGHEGCVATKPGIYTLTKKYQVTKSNL 258

RESULT 7  
AAU82754  
ID AAU82754 standard; Protein; 284 AA.

XX AAU82754;

XX 23-APR-2002 (first entry)

XX Amino acid sequence of novel human protease #53.

XX Human; protease; cancer; immune-related disorder; cardiovascular disease;  
 KW neuronal-associated disease; metabolic disorder; inflammatory disorder;  
 KW nervous system disorder; sexual dysfunction; pain; mood disorder;  
 KW hypertension; psychotic disorder; neurological disorder; dyskinesia;  
 KW viral infection; human immunodeficiency virus; HIV; non-viral infection;  
 KW ocular disease; cytostatic; enzyme.  
 XX Homo sapiens.  
 OS  
 XX WO200200860-A2.  
 PN  
 XX  
 PD 03-JAN-2002.  
 XX  
 XX 26-JUN-2001; 2001WO-US20171.  
 PF  
 XX 26-JUN-2000; 2000US-214047P.  
 PR  
 XX (SUGB-) SUGEN INC.  
 PA  
 ... Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;  
 Charyczak G;  
 DR WPI; 2002-139913/18.  
 DR N-PSDB; ABK31796.  
 XX Nucleic acids encoding novel human proteases, useful for useful for  
 PT treating diseases and disorders such as cancers, immune-related  
 PT diseases and disorders, cardiovascular disease (e.g. restenosis) and  
 PT inflammatory disorders -  
 XX  
 PS Claim 6; Fig 2R; 313pp; English.  
 XX  
 CC The present invention relates to the isolation of novel human  
 CC proteases, and the nucleic acids encoding them. The sequences of  
 CC the invention are useful for treating diseases and disorders such as  
 CC cancers (e.g. breast, colon, lung), immune-related diseases and disorders  
 CC (e.g. inflammatory diseases and asthma), cardiovascular diseases  
 CC (e.g. restenosis and coronary thrombosis), brain or neuronal-associated  
 CC disorders, metabolic disorders (e.g. diabetes, obesity), inflammatory  
 CC peripheral nervous system diseases, migraines, pain, sexual dysfunction,  
 CC mood disorders, attention disorders, cognition disorders, hypotension,  
 CC hypertension, psychotic disorders, neurological disorders  
 CC (e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.  
 CC The nucleic acids and polypeptides are also useful for treating viral  
 CC infections caused by human immunodeficiency virus (HIV), and non-viral  
 CC degeneration. AAU82702-AAU82760 represent the novel human proteases of  
 CC the invention.  
 XX  
 SQ Query Match 37.2%; Score 451; DB 23; Length 284;  
 Best Local Similarity 42.7%; Pred. No. 1.7e-34;  
 Matches 100; Conservative 34; Mismatches 88; Indels 12; Gaps 6;  
 QY 1 ILGGEAEAHARYMASVOLNGAHLCGGVLAQWVLSAAHCL--EDAADGKQVVLGAH 58  
 DB 35 IIGGHEVTPHSPRYMASVRFGGHGGGFLLRARVWVSAAHCFSHRDLRTG--LVVLGAH 92  
 QY 59 SLSOPEPSKRLVDLRAVPHPSQPTIDHDLLELQLSEKATLGPVRLPWQRVDRD 116  
 DB 93 VLSTAEPQOVFGIDALTTHPDYHPMTHANDICLLQNGSVAVLGPVGLLRPGRRA-RP 151  
 QY 117 VAPGTCDDVAGMGINVHAGRRPDSLOHVLPLVLDRTACNRTHHDGAIATERLMC---AES 173  
 DB 152 PYATGTCRVAGVGFSDFEELPGLMEAKVRVLDPPVCN--SSWKGHILTLCTRSGS 209  
 QY 174 NRRDSCKDGGPLVCGGVLEGGVYTSGRVCGNRKPKGIYTRVASYAANDSVL 227  
 DB 210 HRRGFCSDSGPLVCRNRAHGLVSPGLWCGLDPKTPDYVTQVSAPVAMWDV 263

## RESULT 8

AA055758  
 ID AAR55758 standard; Protein; 232 AA.  
 XX  
 AC AAR55758;  
 XX  
 DT 17-NOV-1994 (first entry)  
 XX  
 DE Serine protease Met-ase.  
 XX  
 KW Serine protease; RNK Met-1; cytolytic granule; leukemia;  
 KW large granular lymphocyte; RNK-16; Met-ase; tumor; diagnosis;  
 KW DNA probe; hybridization; Lopez.  
 XX  
 OS Homo sapiens.  
 PN WO9412647-A.  
 XX  
 PD 09-JUN-1994.  
 XX  
 PF 03-DEC-1993; 93WO-US11736.  
 XX  
 PR 03-DEC-1992; 92US-0990301.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Henderson LE, Powers JC, Sayer T, Smyth MJ, Sowder R;  
 PI Wiltstrout TA;  
 XX  
 DR WPI; 1994-200275/24.  
 DR N-PSDB; AA066910.  
 XX  
 PT New serine protease with Met-ase activity - isolated from  
 PT lymphocyte granules, also related DNA vectors, transformed cells  
 PT and antibodies.  
 XX  
 PS Disclosure; Page 38-39; 57pp; English.  
 XX  
 CC A new 30 kDa serine protease, RNK Met-1, was isolated from the  
 CC cytolytic granules of rat RNK-16 large granular lymphocyte leukemia  
 CC cells. cDNA was obtained from a rat RNK-16 lambda-gt11 library and  
 CC sequenced (AA066909). The mature protein has the sequence given in  
 CC AAR55757. The corresponding sequences for the human serine protease,  
 CC Met-ase, obtained from Lopez large granular lymphocyte leukemia  
 CC cells are given in AA066910 and AAR55758. DNA encoding RNK Met-1  
 CC can be used as a probe to detect the enzyme in tumor biopsy samples.  
 XX  
 SQ Sequence 232 AA;  
 Query Match 37.0%; Score 447.5; DB 15; Length 232;  
 Best Local Similarity 42.4%; Pred. No. 2.9e-34;  
 Matches 98; Conservative 38; Mismatches 86; Indels 9; Gaps 5;  
 QY 1 ILGGEAEAHARYMASVOLNGAHLCGGVLAQWVLSAAHCL--EDAADGKQVVLGAH 60  
 DB 35 IIGGHEVTPHSPRYMASVRFGGHGGGFLLRARVWVSAAHCFSHRDLRTG--LVVLGAH 92  
 QY 59 SLSOPEPSKRLVDLRAVPHPSQPTIDHDLLELQLSEKATLGPVRLPWQRVDRD 116  
 DB 93 VLSTAEPQOVFGIDALTTHPDYHPMTHANDICLLQNGSVAVLGPVGLLRPGRRA-RP 151  
 QY 117 VAPGTCDDVAGMGINVHAGRRPDSLOHVLPLVLDRTACNRTHHDGAIATERLMC---AES 173  
 DB 152 PYATGTCRVAGVGFSDFEELPGLMEAKVRVLDPPVCN--SSWKGHILTLCTRSGS 209  
 QY 174 NRRDSCKDGGPLVCGGVLEGGVYTSGRVCGNRKPKGIYTRVASYAANDSVL 227  
 DB 210 HRRGFCSDSGPLVCRNRAHGLVSPGLWCGLDPKTPDYVTQVSAPVAMWDV 263

## RESULT 9

AA28640

ID XX AAY28640 standard; Protein; 283 AA.  
 AC AAY28640;  
 XX 03-NOV-1999 (first entry)  
 XX Human secreted protein from cDNA clone HMWJH67.  
 DE Human secreted protein; human cDNA clone HMWJH67; serine protease family;  
 KW putative preproadipsin; serine protease inhibitor; serpin; IgG domain;  
 KW IgG-1; IgG-3; albumin; extracellular matrix degradation; fusion protein;  
 KW His-tag; cancer; arthritis; cardiovascular disorder; tissue regeneration;  
 KW immune system disorder; nervous system disorder.  
 KW Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..31  
 FT /label= Signal\_peptide  
 FT Protein 32..283  
 F /label= Mature\_secreted\_protein  
 X  
 PN W09940183-A1.  
 XX  
 XX 12-AUG-1999. 99WO-US02292.  
 XX  
 XX 04-FEB-1999; 99WO-US02292.  
 XX  
 XX 06-FEB-1998; 98US-0073961.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX N1 J, Ruben SM;  
 PI  
 XX  
 XX WPT; 1999-508502/42.  
 DR N-PSDB; AAX80904.  
 DR  
 XX  
 XX New isolated human serine protease and serpin polypeptides, used to  
 PT develop products for treating e.g. immune disorders, cancers,  
 PT inflammation, transplant rejection or infections, or as food  
 PT additives  
 XX  
 PS Claim 11; Pages 76-77; 99pp; English.  
 XX  
 CC The present sequence is a secreted protein from cDNA clone HMWJH67  
 CC which is isolated from a cDNA library derived from the human bone  
 CC marrow cell line RS4/11. The protein is a member of serine protease  
 CC family and shows a high degree of sequence similarity to putative  
 C- Preproadipsin from Sus scrofa domestica. The protein, its coding sequence  
 C and its inhibitors are used in the diagnosis and treatment of disorders  
 CC related to abnormal level of the protein or mutation in the nucleotide  
 CC sequence. The serine protease is used to develop antagonists and fusion  
 CC proteins. The fusion of this protein to His-tag, HA-tag, IgG domains,  
 CC etc. facilitates purification, fusion to IgG-1, IgG-3 and albumin  
 CC increases the half life time in vivo, etc. The inhibitors and antagonists  
 CC of serine protease can be used for treating disorders characterised by  
 CC degradation of extracellular matrix, e.g. cancer, arthritis, disorders of  
 CC cardiovascular system, immune system, nervous system, etc. The protein  
 CC can also be used to differentiate, proliferate and attract cells leading  
 CC to regeneration of tissues.  
 XX  
 XX Sequence 283 AA;  
 SQ  
 Query Match 36.9%; Score 447; DB 20; Length 283;  
 Best Local Similarity 42.3%; Pred. No. 4e-34;  
 Matches 99; Conservative 35; Mismatches 88; Indels 12; Gaps 6;  
 QY 1 ILGREAHAHAPRYMASVOLNAGLCGVVAEQWLSAAHCL--EDAADGKVOVLGAH 58  
 |::| 1::|::|::|::| 1::|::|::|::| 1::|::|::|::| 1::|::|::|::|  
 Db 34 IIGGEVTPHSPRYMASVRFSGOHGCGFLLRARVWSAAHCFSHRDRTG--LWVLGAH 91  
 |::| 1::|::|::|::| 1::|::|::|::| 1::|::|::|::| 1::|::|::|::|  
 QY 59 SLSQPEPSKRYLVLRVAFPHSDPSQDTIDHDLQLLQLSEKATLGPV--RPLPWQRVDRD 116  
 ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db 92 VLSTAEPTQOVFGIDALTHFDYHPMTWTHANDICLLRLNGSAVLGPAVGLLLRPGRR-RP 150  
 QY 117 VAPGTLCDVAGWGIYVNHAGRRPDSLQHVLLPVLDRATCNRRTTHDGAITERLMC---AES 173  
 ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
 Db 151 PTAGTRCRVAGWGFVDFEELPPGLMEAKVRVLDPDVCN--SSMKGHLTLTLMICTRSGDS 208  
 QY 174 NRDSCKGDGSGPLVCGVLEGVVTSGRVCGNRKKPKGIYTRVASYAAWIDSVL 227  
 ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
 Db 209 HRRGFCSDSGPLVCRNRRAHGLVSFSGLCMGDPKPTPDVYTVQVSFAVAMWDV 262  
 ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
 RESULT 10  
 AAY99371  
 ID AAY99371 standard; Protein; 283 AA.  
 XX  
 XX AAY99371;  
 AC  
 XX  
 DT 08-AUG-2000 (first entry)  
 XX  
 DE Human PRO1549 (UNQ782) amino acid sequence SEQ ID NO:111.  
 XX  
 KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;  
 KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening.  
 KW Homo sapiens.  
 OS  
 XX  
 PN W0200012708-A2.  
 XX  
 XX 09-MAR-2000.  
 PD  
 XX  
 PF 01-SEP-1999; 99WO-US20111.  
 XX  
 XX 01-SEP-1998; 98US-0098716.  
 PR 01-SEP-1998; 98US-0098749.  
 PR 01-SEP-1998; 98US-0098750.  
 PR 02-SEP-1998; 98US-0098803.  
 PR 02-SEP-1998; 98US-0098821.  
 PR 02-SEP-1998; 98US-0098843.  
 PR 09-SEP-1998; 98US-0099536.  
 PR 09-SEP-1998; 98US-0099596.  
 PR 09-SEP-1998; 98US-0099598.  
 PR 09-SEP-1998; 98US-0099602.  
 PR 09-SEP-1998; 98US-0099642.  
 PR 10-SEP-1998; 98US-0099741.  
 PR 10-SEP-1998; 98US-0099754.  
 PR 10-SEP-1998; 98US-0099763.  
 PR 10-SEP-1998; 98US-0099792.  
 PR 10-SEP-1998; 98US-0099808.  
 PR 10-SEP-1998; 98US-0099812.  
 PR 10-SEP-1998; 98US-0099815.  
 PR 10-SEP-1998; 98US-0099816.  
 PR 15-SEP-1998; 98US-0100385.  
 PR 15-SEP-1998; 98US-0100388.  
 PR 15-SEP-1998; 98US-0100390.  
 PR 16-SEP-1998; 98US-0100584.  
 PR 16-SEP-1998; 98US-0100627.  
 PR 16-SEP-1998; 98US-0100661.  
 PR 16-SEP-1998; 98US-0100662.  
 PR 16-SEP-1998; 98US-0100664.  
 PR 17-SEP-1998; 98US-0100683.  
 PR 17-SEP-1998; 98US-0100684.  
 PR 17-SEP-1998; 98US-0100710.  
 PR 17-SEP-1998; 98US-0100711.  
 PR 17-SEP-1998; 98US-0100919.  
 PR 17-SEP-1998; 98US-0100930.  
 PR 18-SEP-1998; 98US-0100848.  
 PR 18-SEP-1998; 98US-0100849.  
 PR 18-SEP-1998; 98US-0101014.  
 PR 18-SEP-1998; 98US-0101058.  
 PR 18-SEP-1998; 98US-0101071.  
 PR 22-SEP-1998; 98US-0101279.  
 PR 23-SEP-1998; 98US-0101471.  
 PR 23-SEP-1998; 98US-0101472.

```

PR 23-SEP-1998; 98US-0101474.
PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101478.
PR 23-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.
PR 29-SEP-1998; 98US-0102207.
PR 29-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102330.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103314.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.
PR 21-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105693.
PR 26-OCT-1998; 98US-0105694.
PR 27-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105881.
PR 27-OCT-1998; 98US-0105882.
PR 27-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106029.
PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106032.
PR 28-OCT-1998; 98US-0106033.
PR 28-OCT-1998; 98US-0106178.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0108500.
PR 30-OCT-1998; 98US-0106454.
PR 03-NOV-1998; 98US-0106856.
PR 03-NOV-1998; 98US-0106902.
PR 03-NOV-1998; 98US-0106905.
PR 03-NOV-1998; 98US-0106919.
PR 03-NOV-1998; 98US-0106932.
PR 03-NOV-1998; 98US-0106934.
PR 10-NOV-1998; 98US-0107783.
PR 17-NOV-1998; 98US-0108775.
PR 17-NOV-1998; 98US-0108779.
PR 17-NOV-1998; 98US-0108787.
PR 17-NOV-1998; 98US-0108788.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108802.
PR 17-NOV-1998; 98US-0108806.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108867.
PR 17-NOV-1998; 98US-0108867.

PR 17-NOV-1998; 98US-0108925.
PR 18-NOV-1998; 98US-0108848.
PR 18-NOV-1998; 98US-0108849.
PR 18-NOV-1998; 98US-0108850.
PR 18-NOV-1998; 98US-0108851.
PR 18-NOV-1998; 98US-0108852.
PR 18-NOV-1998; 98US-0108853.
PR 18-NOV-1998; 98US-0108858.
PR 18-NOV-1998; 98US-0108904.
XX (GETH ) GENENTECH INC.
XX
XX Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
XX WPI: 2000-237871/20.
XX N-PSDB; AAA37053.
XX
XX New mammalian DNA sequences encoding transmembrane, receptor or
XX secreted PRO polypeptides, useful for screening of potential peptide or
XX small molecule inhibitors of the relevant receptor/ligand interactions
XX
XX Claim 12; Fig 64; 773pp; English.
XX
XX AAA37022 to AAA37144 encode the new isolated human transmembrane,
XX receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The
XX transmembrane and receptor PRO proteins can be used for screening of
XX potential peptide or small molecule inhibitors of the relevant
XX receptor/ligand interactions. The polypeptides and nucleotide sequences
XX encoding then have various industrial applications, including uses as
XX pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
XX PCR primers and hybridisation probes used in the isolation of the PRO
XX polypeptides from the present invention.
XX
XX Sequence 283 AA;
XX
XX Query Match 36.9%; Score 447; DB 21; Length 283;
XX Best Local Similarity 42.3%; Pred. No. 4e-34;
XX Matches 99; Conservative 35; Mismatches 88; Indels 12; Gaps 6;
XX
XX QY 1 ILGREAENAHAPYMASVOLNGAHLGGVLAQWVLSAAHCL--EDADGKVVQLLGAH 58
XX :||| | :||||||| :| | | | :| | | | | :| | | | | :| | | | |
XX Db 34 IIGCHEVTPHSRPYMASVRFQGHGCGFLLRARVVSAAHCFSHRDLRTG--LVVLGAH 91
XX :||| | :||||||| :| | | | :| | | | | :| | | | | :| | | | |
XX QY 59 SLSQPEPSKRLRYDLRAVPHDPDSQDPTIDHDLILLQLSEKATLGPV--RPLPWQVDRD 116
XX :||| | :||||||| :| | | | :| | | | | :| | | | | :| | | | |
XX Db 92 VLSTAETQQVFGIDALTPHDYHPMTHANDICLLRLNGSAVLGPAVGLRLPGRRA-RP 150
XX :||| | :||||||| :| | | | :| | | | | :| | | | | :| | | | |
XX QY 117 VAPGTLCDVAGWGIYNHAGRRPDSLSQHVLLPVLDLQSEKATLGPV--RPLPWQVDRD 173
XX :||| | :||||||| :| | | | :| | | | | :| | | | | :| | | | |
XX Db 151 PTAGTRCRVACGFGVDFEELPPGLMEAKRVLDPDVCN--SSWKGHLTLMLCTRSGDS 208
XX :||| | :||||||| :| | | | :| | | | | :| | | | | :| | | | |
XX QY 174 NRRDCKGDSGGPLVCGGVLEGVTSGRVCGNRKKPGIYTRKVASAAYAWDSVL 227
XX :||| | :||||||| :| | | | :| | | | | :| | | | | :| | | | |
XX Db 209 HRRGFCSDSGGPLVCRNRAHGLVFSFGIWCQDPTPDVYTVQVSFAFVAVIWDVV 262
XX :||| | :||||||| :| | | | :| | | | | :| | | | | :| | | | |
XX
XX RESULT 11
XX AAY70026
XX ID AAY70026 standard; Protein; 283 AA.
XX
XX AC AAY70026;
XX
XX XX
XX DT 05-JUN-2000 (first entry)
XX
XX DE Human Protease and associated protein-20 (PPRG-20).
XX
XX Protease and associated protein-20; PPRG-20; anti-PPRG antibody;
XX diagnosis; treatment; cell proliferative disorder; cancer; cirrhosis;
XX arteriosclerosis; atherosclerosis; burstitis; hepatitis; immune disorder;
XX AIDS; Addison's disease; adult respiratory distress syndrome; allergy;
XX ankylosing spondylitis; amyloidosis; cytostatic; antiarteriosclerotic;
XX hepatotrophic; antiinflammatory; virucide; antipsoriatic; anti-HIV;
XX anti-allergic; immunosuppressive; antidiabetic; antianemic;
XX neuroprotective; human.

```



```

PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PA (GETH ) GENENTECH INC.
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WL, Zhang Z;
XX WPI; 2001-602746/68.
XX N-PSDB; AAS46060.
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumours, such as prostate and breast tumours, in mammals and
PT to screen for modulators of the compounds
XX Claim 11; Fig 272; 774pp; English.
XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.
XX SQ Sequence 283 AA;
Query Match 36.9%; Score 447; DB 22; Length 283;
Best Local Similarity 42.3%; Pred. No. 4e-34;
Matches 99; Conservative 35; Mismatches 88; Indels 12; Gaps 6;
OY 1 ILGREAHAHPYMASVQNGALHCGGLVVAEQWLVAASACL--EDAADGKQVVLGAH 58
DB 34 IIGGHEVTPHSRPMYASVREGQHGGLLRARWVYSAACFHRDLRTG--LWVLGAH 91
OY 59 SLQSPESKRLDYLRVAPHPDSQPDTHDILLQLLSEKATLGPV--RPLPQRVDRD 116
DB 92 VLSTAEPTQVFGIDALTTPDYHPMTHANDICLLRLNGSAVLGPAYGLRLPGRRA-RP 150
117 VAPCTLDVAGWGIIVNHAGRRPDSIQHVLLPVLDRATCNRRTHDGAITERLMC---AES 173
DB 151 PTAGTRCRVAGWGFVSDFEELPGLMEAKVRYLDPDVCN--SSWKGHILTLMTICRSQDS 208
OY 174 NRRDSCKGDSGGLPLVCGGVLEGVVTSSRCVGNRRKKGPIYTRVASYAAWIDSVL 227
DB 209 HRRGFCASDGGPLVCRNRAHGLVFSGLWCGDKPTDVTQVSFAVAWIDVV 262
RESULT 13
AAU02497
ID AAU02497 standard; Protein: 283 AA.
AC AAU02497;
XX 07-SEP-2001 (first entry)
XX Human secreted protein TANGO 298.
XX Human secreted protein; TANGO 298; clone jyhMall18f02; chromosome 19p13;
KW bone marrow; complement factor D; alternative complement pathway;
KW complement regulator deficiency; serine protease dysfunction; adipsin;
KW obesity; diabetes; blood and haematopoietic associated disorder;
KW cardiovascular disorder; inflammatory disorder; immune disorder.
XX

```

```

OS Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 1..31
FT /label= Signal_peptide
FT Modified-site 28...33
FT /note= "N-myristoylation site"
FT Protein 32...283
FT /label= Mature_TANGO_298
FT Domain 34...258
FT /label= Trypsin_domain
FT Modified-site 50..52
FT /note= "Protein kinase C phosphorylation site"
FT Modified-site 55...60
FT /note= "N-myristoylation site"
FT Active-site 70..75
FT /note= "Histidine active site common to serine
FT protease, trypsin family"
FT Modified-site 77..79
FT /note= "Protein kinase C phosphorylation site"
FT Modified-site 77..80
FT /note= "Casein kinase II phosphorylation site"
FT Modified-site 94..97
FT /note= "Casein kinase II phosphorylation site"
FT Modified-site 110..113
FT /note= "Casein kinase II phosphorylation site"
FT Modified-site 129..132
FT /note= "Asn is N-glycosylated"
FT Modified-site 144..147
FT /note= "Amidation site"
FT Modified-site 166..169
FT /note= "Casein kinase II phosphorylation site"
FT Modified-site 174..179
FT /note= "N-myristoylation site"
FT Modified-site 189..192
FT /note= "Asn is N-glycosylated"
FT Modified-site 191..193
FT /note= "Protein kinase C phosphorylation site"
FT Modified-site 208..210
FT /note= "Protein kinase C phosphorylation site"
FT Modified-site 236..241
FT /note= "N-myristoylation site"
FT Modified-site 263..266
FT /note= "cAMP and cGMP dependent protein kinase
FT phosphorylation site"
FT Modified-site 275..277
FT /note= "Protein kinase C phosphorylation site"
XX WO200130831-A1.
XX 03-MAY-2001.
XX 27-OCT-2000; 2000WO-US29797.
XX 27-OCT-1999; 99US-0417796.
XX 17-MAY-2000; 2000US-0572275.
XX (MILL-) MILLENNIUM PHARM INC.
XX Fraser CC, Hodge MR;
XX WPI; 2001-300479/31.
XX N-PSDB; AAS04269.
XX New nucleic acid molecule encoding type II transmembrane proteins
XX useful for treating immune related disorders
XX Claim 23; Fig 10; 137pp; English.
XX The present sequence representing a novel human secreted protein
CC TANGO 298 is isolated from clone jyhMall18f02 from a human bone marrow
CC cDNA library. The gene for TANGO 298 maps to chromosome 19p13.3.
CC TANGO 298 shows sequence homology to human adipsin (complement

```





PT helper sequence -

XX PS Disclosure: Fig 6; 67pp; German.

XX  
CC This invention describes a novel method for producing (M1) one or more  
CC serine proteases and/or one or more domains of one or more serine  
CC proteases in a prokaryotic host. (M1) comprises: (1) a serine protease  
CC and/or one or more domains of them are expressed comprising an N-terminal  
CC helper sequence, that comprises a dipeptide with a dipeptide sequence  
CC Met-Y; (2) renaturing the recombinant proteins; and (3) cleaving the  
CC N-terminal sequence. Y = any amino acid, especially proline. The  
CC invention is useful for recombinantly producing serine proteases for the  
CC potential use in the treatment of inflammatory and autoimmune disorders.  
CC The recombinantly produced active proteins comprises only the  
CC catalytically active domain, giving them a specific function with natural  
CC or artificial specificity.

XX SQ Sequence 283 AA;

Very Match 36.8%; Score 446; DB 22; Length 283;

Best Local Similarity 42.3%; Pred. No. 5e-34;

Matches 99; Conservative 35; Mismatches 88; Indels 12; Gaps 6;

QY 1 ILGREAHAHPYMASVOLNGAHLGGVLAQWVLSAAHCL--EDAADGKVQVLLGAH 58

Db 34 IIGHEVTPHSRPMASVRFGGHCGGELLRARVWVSAAHCFSHEDLRTG--LVVLGAH 91

QY 59 SLSQPEPSKRLVDVLRVPHDPDSQPTIDHDLQLLSEKATLGPAV--RPLPWQRVDRD 116

Db 92 VLSTAEPQQVFGIDALTTHPDYHPMTHANDICLLRLNGSAVLGPAVGLLPLPGRRA-RP 150

QY 117 VAPGTLCDVAGWGIYNHAGRRPDSLQHLVLLPVLDRATCNRRTHDGAITERLMC---AES 173

Db 151 PTVGTRCRVAGWGFVSDFEELPPGLMEAKVRVLDPPVCN--SSWKGHLLTLMCTRSQDS 208

QY 174 NRRDCKDGGGLVGGVLEGVVTSGSRVCGNRKPKGIYTRVASYAAMIDSVL 227

Db 209 HRRGFCSDSGGLVLCRNRAHGLVSGGLWCGDKPTPDVYTVQVSAPVAMWDVV 262

Search completed: May 1, 2003, 22:07:13

Job time : 79 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 22:04:21 ; Search time 19 Seconds  
(without alignments)  
1153.613 Million cell updates/sec

Title: US-09-821-255-2

Perfect score: 1211

Sequence: 1 ILGGRERAHRYMASVQL.....KPGIYTRVASYAAMIDSVLA 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1201	99.2	246	1	DBHU	complement factor
2	869.5	71.8	263	1	S5608	complement factor
3	868	71.7	191	2	S54115	complement factor
4	844.5	69.7	259	1	WMMS28	complement factor
5	469.5	38.8	249	2	A55634	granzyme M (EC 3.4
6	456.5	37.7	262	1	A31372	granzyme A (EC 3.4
7	456.5	37.7	264	2	S55663	granzyme 3 (EC 3.4
8	451	37.2	258	2	S56220	trypsin 2 - rat
9	445	36.7	258	2	A45161	serine proteinase
10	434.5	35.9	226	2	S59370	duodenase - bovine
11	427	35.3	251	2	T10262	mast cell serine p
12	426.5	35.2	238	2	S31779	trypsin (EC 3.4.21
13	419.5	34.6	260	2	S56559	neutropsin - mouse
14	416.5	34.4	231	1	TRPGTR	trypsin (EC 3.4.21
15	416.5	34.4	257	2	B45061	granzyme A (EC 3.4
16	416.5	34.4	260	2	A45061	granzyme A (EC 3.4
17	416	34.4	247	2	S05494	trypsin (EC 3.4.21
18	415	34.3	261	2	A29586	tissue kallikrein
19	413.5	34.1	247	1	TRDG	trypsin (EC 3.4.21
20	413.5	34.1	261	1	A32297	seminogelase (EC 3
21	413	34.1	231	2	S31778	trypsin (EC 3.4.21
22	411.5	34.0	261	2	S40162	cathepsin G (EC 3
23	409.5	33.8	243	2	A35871	trypsin (EC 3.4.21
24	409.5	33.8	247	1	A25852	trypsin (EC 3.4.21
25	406	33.5	246	2	QJ1472	trypsin (EC 3.4.21
26	406	33.5	246	2	QJ1471	trypsin (EC 3.4.21
27	406	33.5	249	1	A35842	chymase (EC 3.4.21
28	405	33.4	248	2	A43520	natural killer cel
29	404	33.4	239	2	A27207	tissue kallikrein

30	404	33.4	242	2	S31775	trypsin (EC 3.4.21
31	404	33.4	242	2	S31776	trypsin (EC 3.4.21
32	404	33.4	248	2	S43259	granzyme-like prot
33	403	33.3	247	1	KVHUCM	chymase (EC 3.4.21
34	401.5	33.2	304	2	S33496	trypsin (EC 3.4.21
35	399.5	33.0	247	2	A27547	trypsin (EC 3.4.21
36	399	32.9	226	2	JE0151	myonase (EC 3.4.-
37	399	32.9	246	2	B38678	mast cell proteinase
38	399	32.9	259	2	A29746	tissue kallikrein
39	398.5	32.9	246	2	B25528	trypsin (EC 3.4.21
40	398	32.9	244	2	S26042	chymase (EC 3.4.21
41	397.5	32.8	246	1	TRDGS	trypsin (EC 3.4.21
42	397	32.8	229	1	TRDFS	trypsin (EC 3.4.21
43	397	32.8	242	2	S49489	trypsin (EC 3.4.21
44	397	32.8	246	2	A32692	cytotoxic T-lympho
45	395.5	32.7	229	1	TRBOTR	trypsin (EC 3.4.21

#### ALIGNMENTS

##### RESULT 1

DBHU

complement factor D (EC 3.4.21.46) precursor [validated] - human (fragment)  
N:Alternate names: adipsin; C3 convertase activator  
C:Species: Homo sapiens (man)  
C:Date: 28-Aug-1985 #sequence\_revision 31-Dec-1992 #text\_change 08-Dec-2000  
C:Accession: A40197; A00936; A60571; S66645  
R:White, R.T.; Damm, D.; Hancock, N.; Rosen, B.S.; Lowell, B.B.; Usher, P.; Flier, J  
J. Biol. Chem. 267, 9210-9213, 1992  
A:Title: Human adipsin is identical to complement factor D and is expressed at high  
A:Reference number: A40197; MUID:92250520; PMID:1374388  
A:Accession: A40197  
A:Molecule type: mRNA  
A:Residues: 1-246 <WHI>  
A:Cross-references: CB:M84526  
R:Niemann, M.A.; Bhow, A.S.; Bennett, J.C.; Volanakis, J.E.  
Biochemistry 23, 2482-2486, 1984  
A:Title: Amino acid sequence of human D of the alternative complement pathway.  
A:Reference number: A00936; MUID:85000441; PMID:6383466  
A:Accession: A00936  
A:Molecule type: protein  
A:Residues: 19-44 'G', 46-51 'Q', 53-75 'TH', 78 'P', 80-83 'XXXITIE', 90-172, 86-91, 185-2  
A:Note: a few residues were assigned from the previously published sequence of Reid  
R:Miyata, T.; Oda, O.; Inagi, R.; Sugiyama, S.; Miyama, A.; Maeda, K.; Nakashima, I.  
Mol. Immunol. 27, 637-644, 1990  
A:Title: Molecular and functional identification and purification of complement comp  
A:Reference number: A60571; MUID:90370044; PMID:2395435  
A:Accession: A60571  
A:Molecule type: protein  
A:Residues: 19-20, 'XX', 23-27, 'XX', 30-31, 'XX', 34, 'X', 36-40 <MIY>  
R:Balke, N.; Holtkamp, U.; Hoerl, W.H.; Tschesche, H.  
FEBS Lett. 371, 300-302, 1995  
A:Title: Inhibition of degranulation of human polymorphonuclear leukocytes by comple  
A:Reference number: S66645; MUID:96013156; PMID:7556615  
A:Accession: S66645  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 19-44 'C', 46-48 <BAL>  
C:Comment: Factor D cleaves factor B when the latter is complexed with factor C3b, a  
C:Genetics:  
A:Gene: GDB:DF  
A:Cross-references: GDB:132645; OMIM:134350  
A:Map position: Xpter-Xqter  
C:Superfamily: trypsin; trypsin homology.  
C:Keywords: complement alternate pathway; hydrolase; plasma; serine proteinase  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-246/Product: complement factor D (fragment) #status experimental <MAT>  
F:19-241/Domain: trypsin homology <TRY>  
F:44-60, 141-207, 172-188, 197-222/Disulfide bonds: #status predicted  
F:59, 105, 201/Active site: His, Asp, Ser #status predicted

Query Match 99.2% Score 1201; DB 1; Length 246;

Best Local Similarity 99.1%; Pred. No. 9.5e-102;  
Matches 226; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILGGREAEAHARYWASVOLNCAHLGCGVLAEOVLSAAHLEDAADGKVQVLLGAHSL 60  
|||||  
DB 19 ILGGREAEAHARYWASVOLNCAHLGAGVLAERWVLSAAHLEDAADGKVQVLLGAHSL 78  
|||||

QY 61 SQPEPSKRLDYDLRAVPHDPDSQPTIDHLLQLLSEKATLGPVAVRPLPWQVRDRAVAPG 120  
|||||  
DB 79 SQPEPSKRLDYDLRAVPHDPDSQPTIDHLLQLLSEKATLGPVAVRPLPWQVRDRAVAPG 138  
|||||

QY 121 TLCDVAGMGIVNHAGRRPDSLOHVLPLVLDRAVPHDPDSQPTIDHLLQLLSEKATLGPVAVRPLPWQVRDRAVAPG 180  
|||||  
DB 139 TLCDVAGMGIVNHAGRRPDSLOHVLPLVLDRAVPHDPDSQPTIDHLLQLLSEKATLGPVAVRPLPWQVRDRAVAPG 198  
|||||

QY 181 GDSGGPLVCGVLEGVTSGRVCGNRRKPGIYTRVASYAAMIDSVLA 228  
|||||  
DB 199 GDSGGPLVCGVLEGVTSGRVCGNRRKPGIYTRVASYAAMIDSVLA 246  
|||||

F T 2  
1 8  
Complement factor D (EC 3.4.21.46) precursor - rat  
N:Alternate names: adipsin; C3 convertase activator; endogenous vascular elastase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 19-May-2000  
C:Accession: S19275  
R:Zhu, L.; Wigle, D.; Hinek, A.; Kobayashi, J.; Ye, C.; Zuker, M.; Dodo, H.; Keeley, F.W.  
J. Clin. Invest. 94, 1163-1171, 1994  
A:Title: The endogenous vascular elastase that governs development and progression of m  
see comments.  
A:Reference number: I55608; MUID:94365184; PMID:8083356  
A:Accession: I55608  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-263 <RES>  
A:Cross-references: GB:S73894; NID:9693721; PIDN:AAB31922.1; PID:g693722  
A:Experimental source: pulmonary artery C.J.; Turcatti, G.  
R:Baker, B.C.; Campbell, C.J.; Grinham, C.J.; Turcatti, G.  
Biochem. J. 279, 775-779, 1991  
A:Title: Purification and partial characterization of rat factor D.  
F:46,124,256,260/Binding site: carboxylate bonds: #status predicted  
F:51-67,149-215,180-205,230-230/Disulfide bonds: #status predicted  
F:66,115,209/Active site: His, Asp, Ser #status predicted

Query Match 71.8%; Score 869.5; DB 1; Length 263;  
Best Local Similarity 69.0%; Pred. No. 1.5e-71;  
Matches 158; Conservative 31; Mismatches 39; Indels 1; Gaps 1;

QY 1 ILGGREAEAHARYWASVOLNCAHLGCGVLAEOVLSAAHLEDAADGKVQVLLGAHSL 59  
|||||  
DB 26 ILGGREAEAHARYWASVOLNCAHLGCGVLAEOVLSAAHLEDAADGKVQVLLGAHSL 85  
|||||

QY 60 LSQPEPSKRLDYDLRAVPHDPDSQPTIDHLLQLLSEKATLGPVAVRPLPWQVRDRAVAPG 119  
|||||  
DB 86 LSQPEPSKRLDYDLRAVPHDPDSQPTIDHLLQLLSEKATLGPVAVRPLPWQVRDRAVAPG 145  
|||||

QY 130 GTLCDVAGMGIVNHAGRRPDSLOHVLPLVLDRAVPHDPDSQPTIDHLLQLLSEKATLGPVAVRPLPWQVRDRAVAPG 179  
|||||  
DB 146 GTLCDVAGMGIVNHAGRRPDSLOHVLPLVLDRAVPHDPDSQPTIDHLLQLLSEKATLGPVAVRPLPWQVRDRAVAPG 205  
|||||

QY 180 KDSGGPLVCGVLEGVTSGRVCGNRRKPGIYTRVASYAAMIDSVLA 228  
|||||  
DB 206 KDSGGPLVCGVLEGVTSGRVCGNRRKPGIYTRVASYAAMIDSVLA 254  
|||||

RESULT 3  
S54115  
complement factor D (EC 3.4.21.46) - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 19-May-2000  
C:Accession: S54115  
R:Nicolas, N.  
submitted to the EMBL Data Library, April 1995  
A:Reference number: S54115  
A:Accession: S54115  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-191 <NIC>  
A:Cross-references: EMBL:Z49058; NID:g773264; PIDN:CAA8844.1; PID:g773265  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: complement alternate pathway; hydrolase; plasma; serine proteinase  
F:1-181/Domain: trypsin homology (fragment) <TRY>

Query Match 71.7%; Score 868; DB 2; Length 191;  
Best Local Similarity 85.5%; Pred. No. 1.4e-71;  
Matches 159; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

QY 43 LEDAAGKQVQLLGAHSLSQPEPSKRLDYDLRAVPHDPDSQPTIDHLLQLLSEKATLGPVAVRPLPWQVRDRAVAPG 102  
|||||  
DB 1 LEDAAGKQVQLLGAHSLSQPEPSKRLDYDLRAVPHDPDSQPTIDHLLQLLSEKATLGPVAVRPLPWQVRDRAVAPG 60  
|||||

QY 103 PAVRPLPWQVRDRAVAPGTLCDVAGMGIVNHAGRRPDSLOHVLPLVLDRAVPHDPDSQPTIDHLLQLLSEKATLGPVAVRPLPWQVRDRAVAPG 162  
|||||  
DB 61 PAVRPLPWQVRDRAVAPGTLCDVAGMGIVNHAGRRPDSLOHVLPLVLDRAVPHDPDSQPTIDHLLQLLSEKATLGPVAVRPLPWQVRDRAVAPG 120  
|||||

QY 163 ALTERLMCAESNRDSCGSGPLVCGVLEGVTSGRVCGNRRKPGIYTRVASYAAMIDSVLA 222  
|||||  
DB 121 ALTERLMCAESNRDSCGSGPLVCGVLEGVTSGRVCGNRRKPGIYTRVASYAAMIDSVLA 180  
|||||

QY 223 IDSVLA 228  
|||||  
DB 181 IDGVMA 186  
|||||

RESULT 4  
WMS28  
complement factor D (EC 3.4.21.46) precursor - mouse  
N:Alternate names: adipocyte 28K proteinase; adipsin; C3 convertase activator; comple  
C:Species: Mus musculus (house mouse)  
C:Date: 13-Aug-1986 #sequence\_revision 13-Aug-1986 #text\_change 19-May-2000  
C:Accession: C25952; A00937; A26105  
R:Phillips, M.; Djian, P.; Green, H.  
J. Biol. Chem. 261, 10821-10827, 1986  
A:Title: The nucleotide sequence of three genes participating in the adipose differen  
A:Reference number: A92553; MUID:86278164; PMID:3015943  
A:Accession: C25952  
A:Molecule type: cDNA  
A:Residues: 1-259 <PHI>  
A:Cross-references: GB:M13386; NID:g192033; PIDN:AAA37262.1; PID:g387105  
R:Cook, K.S.; Groves, D.L.; Min, H.Y.; Spiegelman, B.M.  
Proc. Natl. Acad. Sci. U.S.A. 82, 6480-6484, 1985  
A:Title: A developmentally regulated mRNA from 3T3 adipocytes encodes a novel serine  
A:Reference number: A00937; MUID:86016726; PMID:3901003  
A:Accession: A00937  
A:Molecule type: mRNA  
A:Residues: 1-259 <COO>  
A:Cross-references: GB:M11768; NID:g202166; PIDN:AAA40486.1; PID:g202167  
A:Experimental source: strain Swiss White  
A:Note: only one Ala is present in place of Ala-19 and Ala-20 in another equally abun  
R:Min, H.Y.; Spiegelman, B.M.  
Nucleic Acids Res. 14, 8879-8892, 1986  
A:Title: Adipsin, the adipocyte serine protease: gene structure and control of expres  
A:Reference number: A26105; MUID:87066764; PMID:3024123

A:Accession: A26105  
A:Molecule type: mRNA  
A:Residues: 1-259 <MIN>  
A:Cross-references: GB:X04673; NID:g49883; PIDN:CAA28378.1; PID:g581866  
C:Comment: Human complement factor D is synthesized primarily in cells of the macrophage system activation. However, expression of the murine homolog is specific to adipose tissue.  
C:Genetics:  
A:Gene: 28K  
A:Introns: 19/1; 71/2; 120/3; 206/3  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: adipose tissue; alternative splicing; hydrolase; serine proteinase  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-25/Domain: propeptide #status predicted <PRO>  
F:26-259/Product: adipsin #status predicted <MAT>  
F:26-249/Domain: trypsin homology <TRY>  
F:51-67,149-215,180-196,205-230/Disulfide bonds: #status predicted  
F:66,115,209/Active site: His, Asp, Ser #status predicted

Query Match 69.7%; Score 844.5; DB 1; Length 259;  
Best Local Similarity 66.5%; Pred. No. 2.7e-69;  
Matches 151; Conservative 33; Mismatches 42; Indels 1; Gaps 1;

QY 1 ILGGRGAHARPYMASVOLNGAHLGGVLVAEQWVLSAAHCLDEAA-DGKVQVLLGAHS 59  
Db 26 ILGGQEAHAHARPYMASVQNGTHVCGGTLLDEQWVLSAAHCHMDGVTDSDSVQVLLGAHS 85  
QY 60 LSQPEPSKRLDYLVRAVPHDPSPQDPTIDHLLQLLSEKATLGPAVRPLPQWRDVRDVP 119  
Db 86 LSAPPEYKRWYDVQSVVPHGSRGSDSLEDDLLFLKLSQNASLGPVHVRPLPQVEDKEVEP 145  
QY 120 GTLCDVAGWGIVNIAGRRPDSLQHLVLPVLDRTNCRTHDGAITERLCAESNRDSC 179  
Db 146 GTLCDVAGWGIVYHAGRRPDPVHLQRLVSRINRRTCNLTLYHDGVVTINMCAESNRDTC 205  
QY 180 KGDGGGLVCGVGVLEGVYTSGRVCGNRKKPGIYTRVASYAAWIDSV 226  
Db 206 RGDGSGPLVCGDAVEGYVWGSRVCGNKKPGVYTRVSSYRWIENI 252

RESULT 5  
A55634  
granzyme M (EC 3.4.21.-) precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 24-Feb-1995 #sequence\_revision 24-Feb-1995 #text\_change 08-Sep-1997  
C:Accession: A55634  
R:Pilat, D.; Fink, T.; Obermaier-Skrobanek, B.; Zimmer, M.; Wekerle, H.; Lichter, P.; Jelinek, J.  
"Title: The human Met-ase gene (GZMM): structure, sequence, and close physical linkage reference number: A55634; MUID:95229142; PMID:7713495  
Accession: A55634  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-249 <PI>  
A:Cross-references: GB:L36922  
C:Genetics:  
A:Gene: GDB:GZMM  
A:Cross-references: GDB:387366; OMIM:600311  
A:Map position: 19p13.3-19p13.3  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase  
F:26-249/Domain: trypsin homology <TRY>

Query Match 38.8%; Score 469.5; DB 2; Length 249;  
Best Local Similarity 43.4%; Pred. No. 3.3e-35;  
Matches 99; Conservative 41; Mismatches 79; Indels 9; Gaps 5;

QY 1 ILGGRGAHARPYMASVOLNGAHLGGVLVAEQWVLSAAHCLDEAA-DGKVQVLLGAHS 60  
Db 26 ILGGREVIHPRPYMASLQRGVSHLGGVLVHPKWLTAHCLAQRM-AQLRLVLGLHTL 84  
QY 61 SQPEPSKRLDYLVRAVPHDPSPQDPTIDHLLQLLSEKATLGPAVRPLPQWRDVRDVP 119  
Db 85 DSGPLT---FHTRAAQIOPRYKVPALNDLALLQDGGKVPKSRITRPLALPSKRWVAA 141

RESULT 8  
I56220  
trypsinase 2 - rat  
C:Specimens: Rattus norvegicus (Norway rat)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996  
C:Accession: I56220  
R:Sayers, T.J.; Willtrout, T.A.; Smyth, M.J.; Ottaway, K.S.; Pillaro, A.M.; Sowder, R.; Hel...

## RESULT 10

S69370

duodenase - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 21-Nov-1998  
C:Accession: S69370  
R:Zamolodchikova, T.S.; Vorotyntseva, T.I.; Nazimov, I.V.; Grishina, G.A.  
E:J. Biochem. 227, 873-879, 1995  
A:Title: Duodenase, a new serine protease of unusual specificity from bovine duodenal mu  
A:Reference number: S69370; MUID:95172076; PMID:7867649  
A:Accession: S69370  
A:Molecule type: protein  
A:Residues: 1-226 <ZAM>  
C:Superfamily: trypsin; trypsin homology  
F:1-218/Domain: trypsin homology <TRY>

Query Match 35.9%; Score 434.5; DB 2; Length 226;  
Best Local Similarity 40.9%; Pred. No. 4.5e-32;  
Matches 95; Conservative 44; Mismatches 78; Indels 15; Gaps 6;

1 ILGREAHAHAPYMASV--QLNG-AHLGGVLAQWVLSAAHCLDAADCKVQVLLGA 57  
D 1 IIGGHEAPHSRPMYAFLLFTSGKSHGGGLVREDFVLTAAHCL-----GSINVTLGA 55

Q 58 HSLSOPEPSKRLDYLVRAVPHDPSQPTIDHLLQLSEKATLGPVAPRPLPQWRDROV 117  
D 56 HNIMERERTQVIVRRIPHPDYNDETANDIMLLKTRKADYDKVSPINLPSLAEV 115

Q 118 APGTLCDVAGWGVNHAGRRPDSLOHVLPLVLDRTATNRTHDGAITERLMCA--ESNR 175  
D 116 KPGMNCVAGWGLGNNPSTDKLQGVLEQSEKCIAREKNIPTQ--ICAGDPSKR 173

Q 176 RDSCKGDSGGPLVCGVLEGGVTSRVCGRNKKPGIYTRVASYAAWIDSVL 227  
D 174 RNSFGSDGGPLVCGVLAQGIIVSYGK---NDGTTDPVYTRISFLPWIKRVM 222

RESULT 11

T10262

mast cell serine proteinase (EC 3.4.21.-) 3 - sheep  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T10262  
R:McAleese, S.M.; Pemberton, A.D.; McGrath, M.E.; Huntley, J.F.; Miller, H.R.P.  
Biochem. J. 333, 801-809, 1998  
A:Title: Sheep mast-cell proteinase-1 and -3: cDNA cloning, primary structure and molecu  
A:Reference number: 217004; MUID:98343972; PMID:9677343  
A:Accession: T10262

Status: preliminary; translated from GB/EMBL/DBJ

Molecule type: mRNA

Residues: 1-251 <MCA>

A:Cross-references: EMBL:Y13462; NID:92632147; PIDN:CAA73859.1; PID:g2791548  
A:Experimental source: bone marrow  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase  
F:20-238/Domain: trypsin homology <TRY>

Query Match 35.3%; Score 427; DB 2; Length 251;  
Best Local Similarity 40.9%; Pred. No. 2.5e-31;  
Matches 95; Conservative 43; Mismatches 80; Indels 14; Gaps 6;

1 ILGREAHAHAPYMASVQ--LNG-AHLGGVLAQWVLSAAHCLDAADCKVQVLLGA 57

D 20 IIGGHEAPHSRPMYAFLLFTSGKSHGGGLVREDFVLTAAHCL-----GSINVTLGA 75

Q 58 HSLSOPEPSKRLDYLVRAVPHDPSQPTIDHLLQLSEKATLGPVAPRPLPQWRDROV 117

D 76 HTITDOERTQVIVRRIPHPDYNDETANDIMLLKTRKADYDKVSPINLPSLEKV 135

Q 118 APGTLCDVAGWGVNHAGRRPDSLOHVLPLVLDRTATNRTHDGAITERLMCA--ESNR 175

D 136 KPGMNCVAGWGLGNNPSTDKLQGVLEQSEKCIAREKNIPTQ--ICAGDPSKR 193

Q 176 RDSCKGDSGGPLVCGVLEGGVTSRVCGRNKKPGIYTRVASYAAWIDSVL 227

Db 194 KDSFLDGGPLVCDGVAQGIIVSYGK---DGGTTPNVYTRISFLSWIQRMT 242

RESULT 12

S31779

trypsin (EC 3.4.21.4) III precursor - Atlantic salmon (fragment)  
C:Species: Salmo salar (Atlantic salmon)  
C:Date: 22-Nov-1993 #sequence\_revision 03-Aug-1995 #text\_change 22-Jun-1999  
C:Accession: S66657; S31779  
R:Male, R.; Lorens, J.B.; Smalas, A.O.; Torrisen, K.R.  
Eur. J. Biochem. 232, 677-685, 1995  
A:Title: Molecular cloning and characterization of anionic and cationic variants of  
A:Reference number: S66657; MUID:96035908; PMID:7556223  
A:Accession: S66657  
A:Molecule type: mRNA  
A:Residues: 1-238 <MAL>

A:Cross-references: EMBL:X70074; NID:964387; PIDN:CAA9679.1; PID:g64388  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase  
F:1-7/Domain: signal sequence (fragment) #status predicted <SIG>  
F:8-15/Domain: activation peptide #status predicted <APT>  
F:16-238/Product: trypsin III #status predicted <MAT>  
F:16-231/Domain: trypsin homology <TRY>

F:22-152,40-56,124-225,131-198,163-177,188-212/Disulfide bonds: #status predicted

F:55-99,192/Active site: His, Asp, Ser #status predicted

Query Match 35.2%; Score 426.5; DB 2; Length 238;  
Best Local Similarity 38.2%; Pred. No. 2.6e-31;  
Matches 89; Conservative 52; Mismatches 75; Indels 17; Gaps 8;

Q 1 ILGREAHAHAPYMASVQVLAQWVLSAAHCLDAADCKVQVLLGAHSL 60

Db 16 IGVGVECRKNSASYQASLQ--SGYHFCGSLISSTWVYSAHCKY---SRIQVRLGEHNI 70

Q 61 SOPEPSKRLDYLVRAVPHDPSQPTIDHLLQLSEKATLGPVAPR--PLPQWRDROVA 118

Db 71 AVNEGTEQFIDSVKVTIMHPSYNSRNLNDIMLIKSPASLSYVSTVALP----SSCAS 126

Q 119 PGTLCDVAGWGVNHAGRRPDSLOHVLPLVLDRTATNRTHDGAITERLMCA--ESNR 175

Db 127 SCTRLVSGMGNLSSSSNYPDTLRLCLDLPILSSSSCN--SAYPQIITSNMFCAFMEGG 184

Q 176 RDSCKGDSGGPLVCGVLEGGVTSRVCGRNKKPGIYTRVASYAAWIDSVL 228

Db 185 KDSCKGDSGGPVVNCGLQGVVSWGYG--CAQRNKPQVYTKVCNRSWISSTMS 236

RESULT 13

S6559

neurospilin - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 21-Jul-2000  
C:Accession: I56559  
R:Chen, Z.L.; Yoshida, S.; Kato, K.; Momota, Y.; Suzuki, J.; Tanaka, T.; Ito, J.; Ni  
J. Neurosci. 15, 5088-5097, 1995  
A:Title: Expression and activity-dependent changes of a novel limbic-serine protease  
A:Reference number: I56559; MUID:95348817; PMID:7623137  
A:Accession: I56559  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-260 <RES>

A:Cross-references: GB:D30785; NID:gl648847; PIDN:BAA06451.1; PID:g1020091

C:Superfamily: trypsin; trypsin homology

F:33-252/Domain: trypsin homology <TRY>

Query Match 34.6%; Score 419.5; DB 2; Length 260;  
Best Local Similarity 37.5%; Pred. No. 1.2e-30;  
Matches 87; Conservative 46; Mismatches 86; Indels 13; Gaps 6;

Q 1 ILGREAHAHAPYMASVQVLAQWVLSAAHCLDAADCKVQVLLGAHSL 60

Db 33 ILEGRECIPIHSPQWQAALFQGERLICGGVILVGDWRVLTAAHCKQ----KYSVRLGDHSL 88

QY 61 SOPEPSKRLDYDLRAVPH--DSQPTIDHLLQLSERATLGPVRLPWRQVDRDV 117  
Db 89 QSRDQPEQEIQAQSIQHPYNNSPEDSHDIMLRILSONSLGDKVKPVLANLCPKV 148  
QY 118 APGLCDVAGHGIVNHAGRR-PDSLOHVLPLVDRATCNRRTHDGAITERLMCA-ESNR 175  
Db 149 --GOKCIISGGTIVTSPOENFNTLNCIAEVKIYSONKCEA--YPGKITCGMVCAGSSNG 204  
QY 176 RDSCKGDSGGPLVCGGVLEGVVTSGSRVCGNKKRPGIYTRVASYAAWIDSVL 227  
Db 205 ADTCGDSGGPLVCGMLGLOGITSWGSDCPGKPEKPGVYTKICTYTTWLKKT 256  
RESULT 14  
TRPGTR  
N:Contains: trypsinoren  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 24-Apr-1984 #sequence\_revision 24-Apr-1984 #text\_change 31-Mar-2000  
C #ession: A90641; A90368; A00947  
R #ies, M.; Rovey, M.; Guidoni, A.; Desnuelle, P.  
Biochim. Biophys. Acta 69, 115-129, 1963  
A:Title: Su le trypsinogene et la trypsine de porc.  
A:Reference number: A90641  
A:Accession: A90641  
A:Molecule type: protein  
A:Residues: 1-10 <CHA>  
R:Hermodson, M.A.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.  
Biochemistry 12, 3146-3153, 1973  
A:Title: Determination of the amino acid sequence of porcine trypsin by sequenator analy  
A:Reference number: A90368; MUID:73258692; PMID:4738933  
A:Accession: A90368  
A:Molecule type: protein  
A:Residues: 9-231 <HER>  
A:Note: at position 20, Ile and Val occur alternatively  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; pancreas; polymorphism; protein digestion; serine proteinase; zym  
F:1-231/Product: trypsinogen #status experimental <ZYM>  
F:1-8/Domain: activation peptide #status experimental <APT>  
F:9-231/Product: trypsin #status experimental <MAT>  
F:9-224/Domain: trypsin homology <TRY>  
F:15-145,33-49,117-218,124-191,156-170,181-205/Disulfide bonds: #status predicted  
F:48,92,185/Active site: His, Asp, Ser #status predicted  
F:60,62,65,70/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted  
Query Match 34.4%; Score 416.5; DB 1; Length 231;  
Best Local Similarity 37.2%; Pred. No. 2e-30;  
Mismatches 82; Indels 13; Gaps 7;  
QY 1 ILGGEAFAHAPRYMASVOLNCAHLGGVLAQWVLSAAHCLDEADGKQVLLGAHSL 60  
Db 9 LVGGTCAANSIPYOVSLN-SGSHFCGSLNSQWVSAACHYK-----SRIQVRLGEHNI 63  
QY 61 SOPEPSKRLDYDLRAVPHDPQPTIDHLLQLSERATLGPVRLPWRQVDRDVAPG 120  
Db 64 DYLEGNEQFINAAKIITHPNFNGNTLNDIMLKSPATLNSRVATVSLPR---SCAAG 121  
QY 121 TLDVAGHGIVNHAGRR-PDSLOHVLPLVDRATCNRRTHDGAITERLMCAE--SNRRD 177  
Db 122 TECLISGNGNTKSSGSSYPSSLQCLKAPVLSDDSC--KSSYPGQTGTGNNICVGFLEGGKD 179  
QY 178 SCKGDSGGPLVCGGVLEGVVTSGSRVCGNKKRPGIYTRVASYAAWIDSVL 228  
Db 180 SCGDSGGPVVCGNLOGIGIVSWGYG-CAQKNKPGVYTKVCVNNWIIQOTIA 229  
RESULT 15  
B45061  
granzyme A (EC 3.4.21.78) precursor, splice form 2 - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 19-May-2000  
C:Accession: B45061; A47590

R:Hersberger, R.J.; Gershenfeld, H.K.; Weissman, I.L.; Su, L.  
J. Biol. Chem. 267, 25488-25493, 1992  
A:Title: Genomic organization of the mouse granzyme A gene. Two mRNAs encode the same  
A:Reference number: A45061; MUID:93094270; PMID:1460043  
A:Accession: B45061  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-257 <HER>  
A:Cross-references: GB:L01429; GB:L00631; NID:g1321600; PIDN:AAA99897.1; PID:g1321602  
R:Gershenfeld, H.K.; Weissman, I.L.  
Science 232, 854-858, 1986  
A:Title: Cloning of a cDNA for a T cell-specific serine protease from a cytotoxic T 1  
A:Reference number: A47590; MUID:86208119; PMID:2422755  
A:Accession: A47590  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 12-257 <GER>  
A:Cross-references: GB:M13226  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase  
F:26-249/Domain: trypsin homology <TRY>  
Query Match 34.4%; Score 416.5; DB 2; Length 257;  
Best Local Similarity 37.5%; Pred. No. 2.3e-30;  
Mismatches 87; Conservative 49; Mismatches 87; Indels 9; Gaps 5;  
QY 1 ILGGEAFAHAPRYMASVOLNCAHLGGVLAQWVLSAAHCLDEADGKQVLLGAHSL 60  
Db 26 IIGGDTVVPHSRPYMALLKLSNTICAGALTEKNWVLTAAHC---NVCKRSKFLGAHSI 82  
QY 61 SOPEPSKRLDYDLRAVPHDPQPTIDHLLQLSERATLGPVRLPWRQVDRDVAPG 120  
Db 83 NK-EPEQOILTVKKAFFPCYDEYTRREGDLQVRLKKATVNRNVAILHLPKGDDVKPG 141  
QY 121 TLDVAGHGIVNHAGRRPDSLOHVLPLVDRATCNRRTHD--CAITERLMCAESNR--R 176  
Db 142 TRCRVAGNCRGKNSAPSETLREYNITVIDRKICNDEKHNFHPVIGLNMICAGDLRGGK 201  
QY 177 DSCGDSGGPLVCGGVLEGVVTSGSRVCGNKKRPGIYTRVA-SYAAWIDSVL 227  
Db 202 DSCNGDSGSPLLCDGILLRGITSGGKCGRRRWPVGVTFSLDKHLNWIKITM 253  
Search completed: May 1, 2003, 22:08:51  
Job time : 20 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 22:03:35 ; Search time 25 Seconds  
(without alignments)  
378.264 Million cell updates/sec

Title: US-09-821-255-2

Perfect score: 1211

Sequence: 1 ILGGREAEHAPYMASVOL.....KPGIYTRVASYAAMIDSVLA 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

# number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1211	100.0	253	1	CFAD_HUMAN
2	1056	87.2	259	1	CFAD_PIG
3	869.5	71.8	263	1	CFAD_RAT
4	844.5	69.7	259	1	CFAD_MOUSE
5	471.5	38.9	257	1	GRAM_HUMAN
6	457.5	37.8	263	1	GRAM_MOUSE
7	456.5	37.7	262	1	GRAM_HUMAN
8	456.5	37.7	264	1	GRAM_HUMAN
9	451	37.2	258	1	GRAM_RAT
10	445	36.7	258	1	GRAM_RAT
11	437.5	36.1	260	1	KLK8_HUMAN
12	434.5	35.9	226	1	DDN1_BOVIN
13	433.5	35.8	244	1	KLK6_HUMAN
14	431.5	35.6	260	1	NRPN_RAT
15	427	35.3	251	1	MCT3_SHEEP
16	426.5	35.2	238	1	TRY3_SALSA
17	425	35.1	245	1	MCT1_SHEEP
18	419.5	34.6	260	1	NRPN_MOUSE
19	418	34.5	277	1	KLK2_HUMAN
20	416.5	34.4	231	1	TRYP_PIG
21	416.5	34.4	260	1	GRAM_MOUSE
22	416	34.4	247	1	TRY4_RAT
23	415	34.3	261	1	KLK2_HUMAN
24	413.5	34.1	247	1	TRY2_CANFA
25	413.5	34.1	261	1	KLK3_HUMAN
26	413	34.1	231	1	TRY3_SALSA
27	411.5	34.0	261	1	CATG_MOUSE
28	409.5	33.8	243	1	TRY1_XENLA
29	409.5	33.8	247	1	TRY1_HUMAN
30	409.5	33.8	256	1	KLKE_HUMAN
31	406	33.5	246	1	TRYA_RAT
32	406	33.5	246	1	TRYB_RAT
33	406	33.5	249	1	MCT1_CANFA

34 405 33.4 248 1 NKPL\_RAT  
35 404 33.4 239 1 KLK2\_CAVPO  
36 404 33.4 242 1 TRY1\_SALSA  
37 403 33.3 247 1 MCT1\_HUMAN  
38 401.5 33.2 244 1 TRY2\_XENLA  
39 401.5 33.2 304 1 TRY4\_HUMAN  
40 401 33.1 247 1 MCT1\_PAPHA  
41 399.5 33.0 247 1 TRY3\_RAT  
42 399 32.9 246 1 MCT4\_MOUSE  
43 399 32.9 259 1 KLKM\_MOUSE  
44 398.5 32.9 246 1 TRY2\_MOUSE  
45 398.5 32.9 251 1 KLKE\_HUMAN

PL18291 rattus norv  
PL12323 cavia porce  
P35031 salmo salar  
P23946 homo sapien  
P70059 xenopus lae  
P35030 homo sapien  
P52195 papio hamad  
P08426 rattus norv  
P21812 mus musculu  
P15948 mus musculu  
P07146 mus musculu  
O9p0g3 homo sapien

#### ALIGNMENTS

RESULT 1  
ID CFAD\_HUMAN STANDARD; PRT; 253 AA.  
AC P00746;  
DT 21-JUL-1986 (Rel. 01, Created)  
DF 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Complement factor D precursor (EC 3.4.21.46) (C3 convertase activator)  
DE (Properdin factor D) (Adipsin).  
GN DF.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_Taxid:9606;  
RN [1]  
RP SEQUENCE OF 8-253 FROM N.A.  
RX MEDLINE=92250520; PubMed=1374388;  
RA White R.T., Damm D., Hancock N., Rosen B.S., Lowell B.B., Usher P.,  
RA Elter J.S., Spiegelman B.M.;  
RT "Human adipsin is identical to complement factor D and is expressed  
RT at high levels in adipose tissue.";  
RL J. Biol. Chem. 267:9210-9213(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Elter J.S., Spiegelman B.M., Rosen B.M.;  
RL Patent number WO9006365, 14-JUN-1990.  
RN [3]  
RP SEQUENCE OF 26-252.  
RX MEDLINE=85000441; PubMed=6383466;  
RA Niemann M.A., Bhown A.S., Bennett J.C., Volanakis J.E.;  
RT "Amino acid sequence of human D of the alternative complement  
RT pathway.";  
RL Biochemistry 23:2482-2486(1984).  
RN [4]  
RP PARTIAL SEQUENCE OF 26-252.  
RX MEDLINE=84108950; PubMed=6361333;  
RA Johnson D.M.A., Gagnon J., Reid K.B.M.;  
RT "Amino acid sequence of human factor D of the complement system.  
RT Similarity in sequence between factor D and proteases of non-plasma  
RT origin.";  
RL FEBS Lett. 166:347-351(1984).  
RN [5]  
RP PARTIAL SEQUENCE OF 26-61 AND 194-220.  
RX MEDLINE=84256515; PubMed=6821372;  
RA Johnson D.M.A., Gagnon J., Reid K.B.M.;  
RT "Factor D of the alternative pathway of human complement.  
RT Purification, alignment and N-terminal amino acid sequences of the  
RT major cyanogen bromide fragments, and localization of the serine  
RT residue at the active site.";  
RL Biochem. J. 187:863-874(1980).  
RN [6]  
RP PARTIAL SEQUENCE OF 26-82.  
RX MEDLINE=80145719; PubMed=6987665;  
RA Volanakis J.E., Bhown A.S., Bennett J.C., Mole J.E.;  
RT "Partial amino acid sequence of human factor D:homology with serine  
RT proteases.";

Proc. Natl. Acad. Sci. U.S.A. 77:1116-1119(1980).  
[7] PARTIAL SEQUENCE OF 26-78.  
MEDLINE-81054886; PubMed-6776531;  
Davis A.E. III;  
\*Active site amino acid sequence of human factor D.\*;  
Proc. Natl. Acad. Sci. U.S.A. 77:4938-4942(1980).  
[8]  
X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
MEDLINE-94118317; PubMed-8289289;  
Narayana S.V.L., Carson M., El-Kabani O., Kilpatrick J.M., Moore D.,  
Chen X., Bug C.E., Volanakis J.E., Delucas L.J.;  
\*Structure of human factor D. A complement system protein at 2.0-A  
resolution.\*;  
J. Mol. Biol. 235:695-708(1994).  
[9]  
X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
MEDLINE-96025834; PubMed-7592653;  
Kim S., Narayana S.V., Volanakis J.E.;  
\*Crystal structure of a complement factor D mutant expressing  
enhanced catalytic activity.\*;  
J. Biol. Chem. 270:24399-24405(1995).  
CC -1- FUNCTION: FACTOR D CLEAVES FACTOR B WHEN THE LATTER IS COMPLEXED  
WITH FACTOR C3B, ACTIVATING THE C3BBB COMPLEX, WHICH THEN BECOMES  
THE C3 CONVERTASE OF THE ALTERNATE PATHWAY. ITS FUNCTION IS  
HOMOLOGOUS TO THAT OF C1S IN THE CLASSICAL PATHWAY.  
CC -1- CATALYTIC ACTIVITY: Cleaves component factor B (Arg-|-Lys) when in  
complex with C3b or with cobra venom factor (CVF).  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- CAUTION: IN ADDITION TO THE CONFLICTS SHOWN IN THE FEATURE TABLE,  
REF.3 SEQUENCE HAD A PEPTIDE IN THE WRONG ORDER AND ANOTHER ONE  
MISSING.  
-----  
THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M84526; AAA35527.1; ALT\_INIT.  
DR PIR; A40197; DBHU.  
DR PDB; 1DFP; 25-FEB-98.  
DR PDB; 1DSU; 11-JUL-96.  
DR PDB; 1DSU; 17-AUG-96.  
DR MEROPS; S01.191;  
DR Genew; HGNC:2771; DF.  
DR MIM; 134350;  
L InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Complement alternate pathway; Plasma; Hydrolase; Serine protease;  
KW Signal; Glycoprotein; Signal; 3D-structure.  
FT SIGNAL 1 20 POTENTIAL.  
FT Zymogen; 21 25 ACTIVATION PEPTIDE (POTENTIAL).  
FT CHAIN 26 253 COMPLEMENT FACTOR D.  
FT ACT\_SITE 66 66 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 112 112 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 208 208 CHARGE RELAY SYSTEM.  
FT DISULFID 51 67  
FT DISULFID 148 214  
FT DISULFID 179 195  
FT DISULFID 204 229  
FT CONFLICT 26 26  
FT CONFLICT 35 35  
FT CONFLICT 40 40  
FT CONFLICT 49 49  
I -> M (IN REF. 1).  
H -> F (IN REF. 6).  
M -> V (IN REF. 6).  
H -> E (IN REF. 4 AND 5).

FT CONFLICT 52 52 G -> A (IN REF. 1 AND 2).  
FT CONFLICT 59 59 Q -> R (IN REF. 1 AND 2).  
FT CONFLICT 63 63 S -> T (IN REF. 4).  
FT CONFLICT 73 73 D -> G (IN REF. 4).  
FT CONFLICT 83 86 HSLS -> THLP (IN REF. 3).  
FT CONFLICT 83 84 HS -> ST (IN REF. 4).  
FT CONFLICT 94 95 MISSING (IN REF. 4).  
FT CONFLICT 96 96 D -> E (IN REF. 4).  
FT CONFLICT 136 136 TCNRRTHDGAITE -> KCRLYDVL (IN REF. 4).  
FT CONFLICT 178 191 Q -> G (IN REF. 4).  
FT CONFLICT 243 243 S -> T (IN REF. 3).  
FT CONFLICT 250 250 S -> H (IN REF. 3).  
FT CONFLICT 250 250 MISSING (IN REF. 4).  
SQ SEQUENCE 253 AA; 27004 MW; BD553B70BD55C6AD CRC64;  
Query Match 100.0%; Score 1211; DB 1; Length 253;  
Best Local Similarity 100.0%; Pred. No. 1.9e-107;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ILGREGAEAHARPYMASVOLNGAHLGCGVLVAEQWVLSAAHCLDEADAGKQVLLGAHSL 60  
DB 26 ILGREGAEAHARPYMASVOLNGAHLGCGVLVAEQWVLSAAHCLDEADAGKQVLLGAHSL 85  
QY 61 SQPEPSKRLYDVLRAVPHPSQPTDIHDLQLLQSLSEKATLGPVRLPWVRDVRDAPG 120  
DB 86 SQPEPSKRLYDVLRAVPHPSQPTDIHDLQLLQSLSEKATLGPVRLPWVRDVRDAPG 145  
QY 121 TLCDDVAGMGVNHAGRPDSLSQHVLLPVLDRATCNRRTHDGAITERLWCAESNRRDCK 180  
DB 146 TLCDDVAGMGVNHAGRPDSLSQHVLLPVLDRATCNRRTHDGAITERLWCAESNRRDCK 205  
QY 181 GDSGGPLVCGGVLEGVTSGRVCGNKKPKGIYTRVASYAAMDSVLA 228  
DB 206 GDSGGPLVCGGVLEGVTSGRVCGNKKPKGIYTRVASYAAMDSVLA 253  
RESULT 2.  
CFAD\_PIG STANDARD; PRT; 259 AA.  
AC P51779;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Complement factor D precursor (EC 3.4.21.46) (C3 convertase activator)  
DE (Properdin factor D) (Adipsin).  
GN DF.  
OS Sus scrofa (pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Adipose tissue;  
RA Miner J.L., Hahn K.J., Staten N.R., Baile C.A.;  
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 69-259 FROM N.A.  
RC TISSUE=Adipose tissue;  
RA Nicolas N.;  
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: FACTOR D CLEAVES FACTOR B WHEN THE LATTER IS COMPLEXED  
WITH FACTOR C3B, ACTIVATING THE C3BBB COMPLEX, WHICH THEN BECOMES  
THE C3 CONVERTASE OF THE ALTERNATE PATHWAY. ITS FUNCTION IS  
HOMOLOGOUS TO THAT OF C1S IN THE CLASSICAL PATHWAY (BY  
SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Cleaves component factor B (Arg-|-Lys) when in  
complex with C3b or with cobra venom factor (CVF).  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

EMBL; U29948; AAA73627.1; -
DR DR
EMBL; Z49058; CAA88844.1; -
DR DR
HSP; P00746; IDSU.
DR DR
MEROPS; S01.191; -
DR DR
InterPro; IPR001314; Chymotrypsin.
DR DR
InterPro; IPR001254; Ser.protease_Try.
DR DR
Pfam; PF00089; trypsin; 1.
DR DR
PRINTS; PR00723; CHYMOTRYPSIN.
DR DR
SMART; SM00020; Tryp_SPC; 1.
DR DR
PROSITE; PS50240; TRYPSIN_DOM; 1.
DR DR
PROSITE; PS00134; TRYPSIN_HIS; 1.
DR DR
PROSITE; PS00135; TRYPSIN_SER; 1.
DR DR
KW Complement alternate pathway; Plasma; Hydrolase; Serine protease;
KW zymogen; Glycoprotein; Signal.
KW SIGNAL 1 21
FT PROPEP 22 26 POTENTIAL.
DR CHAIN 27 259 ACTIVATION PEPTIDE (POTENTIAL).
DR ACT_SITE 67 67 COMPLEMENT FACTOR D.
DR ACT_SITE 115 115 CHARGE RELAY SYSTEM.
DR ACT_SITE 209 209 CHARGE RELAY SYSTEM.
DR DISULFID 52 68 CHARGE RELAY SYSTEM.
DR DISULFID 149 215 BY SIMILARITY.
DR DISULFID 180 196 BY SIMILARITY.
DR DISULFID 205 230 BY SIMILARITY.
DR SEQUENCE 259 AA; 27763 MW; 0121AAAE0E34CALED CRC64;
DR SQ

```

CFAD\_MOUSE STANDARD: PRT: 259 AA.

AC P03953; Q61280;

DT 23-OCT-1986 (Rel. 02, Created)

DT 23-OCT-1986 (Rel. 02, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Complement factor D precursor (EC 3.4.21.46) (C3 convertase activator)

DE (Properdin factor D) (Adipsin) (28 kDa protein, adipocyte).

GN DF OR ADN.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=86278164; PubMed=3015943;

RA Phillips M., Djian P., Green H.;

RT "The nucleotide sequence of three genes participating in the adipose differentiation of 3T3 cells.";

RL J. Biol. Chem. 261:10821-10827(1986).

F [2]

I SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).

RA MEDLINE=87066764; PubMed=3024123;

RA Min H.Y., Spiegelman B.M.;

RT "Adipsin, the adipocyte serine protease: gene structure and control of expression by tumor necrosis factor.";

RL Nucleic Acids Res. 14:8879-8892(1986).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Swiss white;

RX MEDLINE=86016726; PubMed=3901003;

RA Cook K.S., Groves D.L., Min H.Y., Spiegelman B.M.;

RT "A developmentally regulated mRNA from 3T3 adipocytes encodes a novel serine protease homologue.";

RL Proc. Natl. Acad. Sci. U.S.A. 82:6480-6484(1985).

CC -!- FUNCTION: FACTOR D CLEAVES FACTOR B WHEN THE LATTER IS COMPLEXED WITH FACTOR C3B, ACTIVATING THE C3BBB COMPLEX, WHICH THEN BECOMES THE C3 CONVERTASE OF THE ALTERNATE PATHWAY. ITS FUNCTION IS HOMOLOGOUS TO THAT OF C1S IN THE CLASSICAL PATHWAY.

CC -!- CATALYTIC ACTIVITY: Cleaves component factor B (Arg-1-Lys) when in complex with C3b or with cobra venom factor (CVF).

CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are produced by alternative splicing.

CC -!- PTM: N-GLYCOSYLATED.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; M11768; AAA0486.1; -

DR EMBL; X04673; CAA28378.1; -

DR EMBL; X04673; CAA28379.1; -

DR EMBL; M13386; AAA37262.1; -

DR PIR; C25952; WMS28.

DR HSP; P00746; IDSU.

DR MEROPS; S01.191; -

DR MGD; MGI:87931; Adn.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR001254; Ser\_protease\_Try.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; Tryp\_Spc; 1.

DR PROSITE; PS00240; TRYPSIN\_DOM; 1.

DR PROSITE; PS00134; TRYPSIN\_HIS; 1.

DR PROSITE; PS00135; TRYPSIN\_SER; 1.

DR Complement alternate pathway; Plasma; Hydrolase; Serine protease;

KW zymogen; Glycoprotein; Alternative splicing; Signal.

RN [2] OR 19 (POTENTIAL).

FT SIGNAL

FT PROPEP 21 25  
FT CHAIN 26 259  
FT ACT\_SITE 66 66  
FT ACT\_SITE 115 115  
FT ACT\_SITE 209 209  
FT DISULFID 51 67  
FT DISULFID 149 196  
FT DISULFID 180 230  
FT DISULFID 205 230  
FT CARBOHYD 46 46  
FT CARBOHYD 124 124  
FT CARBOHYD 176 176  
FT CARBOHYD 251 251  
FT CARBOHYD 256 256  
FT VARSPLIC 20 20  
SQ SEQUENCE 259 AA; 28057 MW; 8C3A952561247DF9 CRC64;

Query Match 69.7%; Score 844.5; DB 1; Length 259;  
Best Local Similarity 66.5%; Pred. No. 9.5e-73;  
Matches 151; Conservative 33; Mismatches 42; Indels 1; Gaps 1;

QY 1 ILGGREAAHARPYMASVOLNGAHLGCGVLVAPOWLSAAHCLDAA-DGKVVQLLGAHS 59  
DB 26 ILGGQEAHARPYMASVQVNGTHVCGTLLDQWLSAAHCLDAA-DGKVVQLLGAHS 85  
QY 60 LSQPPSKRLDYLRVPHDPDQPTIDHLLQLLQSLKATIGPVRPLPQWRDVRDAP 119  
DB 86 LSAPEPKRWYDQSVYVPHGSRPDSLEDDLILFKLSQNASLGPVHPLPQYEDKEVEP 145  
QY 120 GTLCDVAGGVGNHAGRRRPSDLSQHLVLLPVLDRATCRNRTHDCAITERLWCAESNRDSC 179  
DB 146 GTLCDVAGGVGNHAGRRRPSDLSQHLVLLPVLDRATCRNRTHDCAITERLWCAESNRDSC 205  
QY 180 KDGSGPLVCGVLEGVVTSRVCGRNKKPGIYTRVASYAAWIDSV 226  
DB 206 RDSGSLVCGDAVEGVVTSRVCGRNKKPGIYTRVASYAAWIDSV 252

RESULT 5

GRAM\_HUMAN STANDARD; PRT: 257 AA.

AC P51124;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Granzyme M precursor (EC 3.4.21.-) (Met-ASE) (Natural killer cell granular protease) (HU-Met-1) (Met-1 serine protease).  
GN GZMM OR MET1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94065173; PubMed=8245461;  
RA Smyth M.J., Sayers T.J., Wilttrout T., Powers J.C., Trapani J.A.;

RT "Met-ase: cloning and distinct chromosomal location of a serine protease preferentially expressed in human natural killer cells.";

RL J. Immunol. 151:6195-6205(1993).

RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95229142; PubMed=7713495;  
RA Pilat D., Fink T.M., Obermaier-Skrobanek B., Zimmer M., Wekerle H., Lichter P., Jenne D.E.;

RT "The human Met-ase gene (GZMM): structure, sequence, and close physical linkage to the serine protease gene cluster on 19p13.3.";

RL Genomics 24:445-450(1994).

RN [3]  
RP SEQUENCE FROM N.A.  
RA Smyth M.J.;

RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: CLEAVES PEPTIDE SUBSTRATES AFTER METHIONINE, LEUCINE, AND NORLEUCINE.

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 22:03:56 ; Search time 33 Seconds  
(without alignments)  
1423.598 Million cell updates/sec

Title: US-09-821-255-2

Perfect score: 1211

Sequence: 1 ILGREAFAHAPYMASVOL.....KPGYITRVASYAAWIDSVLA 228

Scoring table: BLOSUM62.

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

all number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rotent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1211	100.0	253	4	Q8WZB4	Q8WZB4 homo sapien
2	554	45.7	255	13	O34289	O34289 salvelinus
3	457.5	37.8	256	11	Q9R0K0	Q9R0K0 mus musculus
4	436	36.0	264	11	O08643	O08643 mus musculus
5	432.5	35.7	240	13	Q98TH0	Q98TH0 engraulis j
6	429	35.4	244	13	O42159	O42159 petromyzon
7	429	35.4	245	13	O42160	O42160 petromyzon
8	428	35.3	251	6	Q9GLN2	Q9GLN2 bos taurus
9	425.5	35.1	251	11	O54854	O54854 rattus norv
10	421.5	34.8	246	11	O88301	O88301 mus musculus
11	421.5	34.8	253	11	Q91Y82	Q91Y82 mus musculus
12	419	34.6	247	11	O9CPN7	O9CPN7 mus musculus
13	418	34.5	247	13	O42608	O42608 petromyzon
14	417	34.4	247	13	O42158	O42158 petromyzon
15	414.5	34.2	241	13	Q98TG9	Q98TG9 engraulis j
16	408.5	33.7	255	4	Q96RQ0	Q96RQ0 homo sapien

17	404	33.4	248	11	Q63224	Q63224 rattus norv
18	401	33.1	242	13	Q93266	Q93266 pseudopleur
19	399.5	33.0	247	11	Q9CPN9	Q9CPN9 mus musculus
20	398	32.9	239	11	Q9EOT2	Q9EOT2 mus musculus
21	398	32.9	246	11	Q9EPQ9	Q9EPQ9 mus musculus
22	397	32.8	242	13	Q9WQ77	Q9WQ77 paralicthy
23	397	32.8	242	13	Q92099	Q92099 paranotothe
24	397	32.8	244	13	O8QW33	O8QW33 anguilla ja
25	395.5	32.7	247	11	Q9D7V7	Q9D7V7 mus musculus
26	391.5	32.3	250	13	Q93265	Q93265 pseudopleur
27	391	32.3	261	6	Q9N1O1	Q9N1O1 saguinus oe
28	390.5	32.2	246	11	Q921R9	Q921R9 mus musculus
29	390	32.2	249	13	Q9W6K0	Q9W6K0 notothenia
30	390	32.2	249	13	Q92046	Q92046 dissostichu
31	388.5	32.1	238	13	Q9W7Q6	Q9W7Q6 paralicthy
32	387	32.0	237	13	Q91515	Q91515 fugu rubrip
33	387	32.0	675	13	Q9W6J8	Q9W6J8 dissostichu
34	385.5	31.8	246	11	Q9R0T7	Q9R0T7 mus musculus
35	384	31.7	254	6	Q9XSN6	Q9XSN6 sus scrofa
36	384	31.7	263	11	Q9JM69	Q9JM69 mus musculus
37	383.5	31.7	246	11	Q9QUK9	Q9QUK9 mus musculus
38	383	31.6	344	13	Q9W6J9	Q9W6J9 dissostichu
39	382.5	31.6	247	13	Q9W7Q5	Q9W7Q5 paralicthy
40	382	31.5	234	11	Q9CV76	Q9CV76 mus musculus
41	381.5	31.5	262	4	Q8RCV8	Q8RCV8 homo sapien
42	381	31.5	159	6	Q9GMD8	Q9GMD8 ornithorhyn
43	381	31.5	331	11	O8RIA6	O8RIA6 mus musculus
44	380.5	31.4	254	11	Q61096	Q61096 mus musculus
45	379	31.3	806	6	O18783	O18783 macropus eu

#### ALIGNMENTS

#### RESULT 1

Q8WZB4 PRELIMINARY; PRT; 253 AA.

AC Q8WZB4:

DT 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Adipsin/complement factor D precursor (EC 3.4.21.46).

GN DF.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Relle M.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases;

DR EMBL; AJ313463; CAC48304.1; -

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR001254; Ser\_protease\_Try.

DR Pfam; PF00089; trypsin\_1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; Tryp\_Sec; 1.

DR PROSITE; PS00240; TRYPSIN\_DOM; 1.

DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.

DR PROSITE; PS00135; TRYPSIN\_SER; UNKNOWN\_1.

DR SIGNAL; Hydrolase.

FT SIGNAL 1 8 POTENTIAL.

FT CHAIN 9 253 ADIPSIN/COMPLEMENT FACTOR D PRECURSOR.

SQ SEQUENCE 253 AA; 27033 MW; 78B06C209DEEA362 CRC64;

Query Match 100.0%; Score 1211; DB 4; Length 253;  
Best Local Similarity 100.0%; Pred. No. 5.6e-103;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILGREAFAHAPYMASVOLNGAHLCCGVLVAEQWLSAAHCLDAADAKGVQLGAHSL 60  
|||||  
DB 26 ILGREAFAHAPYMASVOLNGAHLCCGVLVAEQWLSAAHCLDAADAKGVQLGAHSL 85

```
QY 61 SQPEPSKRLYDLVRAVPHPSQPDTHDHLQLLSEKATLGPVRLPWQRVDRDVG 120
Db 86 SQPEPSKRLYDLVRAVPHPSQPDTHDHLQLLSEKATLGPVRLPWQRVDRDVG 145
QY 121 TLCDVAGMGIVNHAGRRPDSLQHVLLPVLDRATCNRRTHHDGAIATERLMAESNRDSC 180
Db 146 TLCDVAGMGIVNHAGRRPDSLQHVLLPVLDRATCNRRTHHDGAIATERLMAESNRDSC 205
QY 181 GDSGGLVCGGVLEGVVTSGRVCCNRRKPGIYTRVSYAAWIDSVLA 228
Db 206 GDSGGLVCGGVLEGVVTSGRVCCNRRKPGIYTRVSYAAWIDSVLA 253

RESULT 2
ID O34289 PRELIMINARY; PRT: 255 AA.
AC O34289;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
F Serine protease-like protein precursor.
C Salvelinus fontinalis (Brook trout) (Brook char).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8038;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PREOVULATORY OVARIAN TISSUE;
RX MEDLINE=98206794; PubMed=9546717;
RA Hajnik C.A., Goetz F.W., Hsu S.Y., Sokal N.;
RT *Characterization of a ribonucleic acid transcript from the brook
RT trout (Salvelinus fontinalis) ovary with structural similarities to
RT mammalian adiponectin/complement factor D and tissue kallikrein, and the
RT effects of kallikrein-like serine proteases on follicle contraction.*;
RL Biol. Reprod. 58:887-897(1998).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: AF005026; AAC1927.1; -.
DR HSSP; P00761; 1EPT.
DR MEROPS; S01.191; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
F Hydrolyase; Protease; Serine protease; Signal.
F SIGNAL 1 22 POTENTIAL.
FT CHAIN 36 255 SERINE PROTEASE-LIKE PROTEIN.
SQ SEQUENCE 255 AA; 27658 MW; 9FB8A0B2C38648FA CRC64;

Query Match 45.7%; Score 554; DB 13; Length 255;
Best Local Similarity 44.8%; Pred. No. 7.7e-43;
Matches 104; Conservative 46; Mismatches 76; Indels 6; Gaps 3;

QY 1 ILGREAFAHARPYMASVOLNGA----HLGGVLVAEQWVLSAAHCLDRAADGKVQVLLG 56
Db 23 ITGREAFAHARPYMASVOLNGA----HLGGVLVAEQWVLSAAHCLDRAADGKVQVLLG 81
QY 57 AHSLSQPEPSKRLYDLVRAVPHPSQPDTHDHLQLLSEKATLGPVRLPWQRVDRDVG-R 115
Db 82 AHSLSQPEPSKRLYDLVRAVPHPSQPDTHDHLQLLSEKATLGPVRLPWQRVDRDVG 141
QY 116 DVAPGTLCDVAGMGIVNHAGRRPDSLQHVLLPVLDRATCNRRTHHDGAIATERLMAESNR 175
Db 142 DPATEQEVNTAGWSLNLGSRDQLQELVIDYNNRVRCGRSDYGGKFTNNMLCAASRQ 201
QY 176 RDSCKGDSGGLVCGGVLEGVVTSGRVCCNRRKPGIYTRVSYAAWIDSVLA 227
Db 202 RDTGDSGGGLYKGVAVGITSNGGKCGSSKPKGLYTTISHYSQWIDKTM 253
```

```
RESULT 3
Q9R0K0
ID Q9R0K0 PRELIMINARY; PRT: 256 AA.
AC Q9R0K0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Granzyme K.
GN GZMK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DDY;
RX MEDLINE=99337786; PubMed=10407175;
RA Suemoto T., Taniguchi M., Shiosaka S., Yoshida S.;
RT *cDNA cloning and expression of a novel serine protease in the mouse
RT brain.*;
RL Brain Res. Mol. Brain Res. 70:273-281(1999).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: AB032200; BAA84221.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.146; -.
DR MGD; MGI:1298232; GzmK.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 256 AA; 28588 MW; AC0CE7418EFD16FB CRC61;

Query Match 37.8%; Score 457.5; DB 11; Length 256;
Best Local Similarity 41.3%; Pred. No. 5.2e-34;
Matches 100; Conservative 42; Mismatches 77; Indels 23; Gaps 9;

QY 1 ILGREAFAHARPYMASVOLNGAHLGGVLVAEQWVLSAAHCLDRAADG-KVQVLLGAHS 59
Db 19 IIGREVOQPHSRPFMASIOYRSKHICGGVLHPQWVLTAAHCYSWFFRGHSPTVVLGAHS 78
QY 60 LSQPEPSKRLYDLVRAVPHPSQPDTHDHLQLLSEKATLGPVRLPWQRVDRDVG 117
Db 79 LSKNEPMKQTEIKKFIFFSRLOSGSASHDMLKRTAAELNNKVVQLHLGSKNYLRD- 137
QY 118 APGTLCDVAGMGIVNHAGRRP-----DSLQHVLLPVLDRATCNRRTH--HDGAIATERL 169
Db 138 --GTRKQVWGTT-----KPDLLTASDTLREVTVTIISKRCSQSYYNKKPVITKDM 190
QY 170 CAESNR--RDSCKGDSGGLVCGGVLEGVVTSGRVCCNRRKPGIYTRVA-SYAAWIDSV 226
Db 191 CAGDARGQKDSGSGGLGCKGIFHALVSQGYK-CGTAKKPGIYTLTKYQTIWIKS 249
QY 227 LA 228
Db 250 LA 251

RESULT 4
O08643
ID O08643 PRELIMINARY; PRT: 264 AA.
AC O08643;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE LYMPOCTYCE MET-ASE 1 precursor (LMET-1).
```

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 22:08:31 ; Search time 46 Seconds  
(without alignments)  
427.682 Million cell updates/sec

Title: US-09-821-255-2  
Perfect score: 1211  
Sequence: 1 ILGREAHAHPYMASVOL.....KPGIYTRVASYAAMIDSVLA 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues

al number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1211	100.0	228	10	US-09-821-255-2
2	1056	87.2	233	10	US-09-821-255-4
3	456.5	37.7	264	9	US-10-114-893-26
4	451	37.2	284	10	US-09-888-615-112
5	447	36.9	283	9	US-10-174-590-272
6	447	36.9	283	9	US-10-176-758-272
7	447	36.9	283	9	US-10-175-737-272
8	447	36.9	283	9	US-10-173-706-272
9	447	36.9	283	9	US-10-175-738-272
10	447	36.9	283	9	US-10-175-752-272
11	447	36.9	283	9	US-10-176-482-272
12	447	36.9	283	9	US-10-176-757-272
13	447	36.9	283	9	US-10-176-913-272
14	447	36.9	283	9	US-10-180-552-272
15	447	36.9	283	9	US-10-180-557-272
16	447	36.9	283	9	US-10-173-700-272
17	447	36.9	283	9	US-10-174-572-272
18	447	36.9	283	9	US-10-174-579-272
19	447	36.9	283	9	US-10-174-582-272

20	447	36.9	283	9	US-10-174-588-272	Sequence 272, App
21	447	36.9	283	9	US-10-175-739-272	Sequence 272, App
22	447	36.9	283	9	US-10-175-740-272	Sequence 272, App
23	447	36.9	283	9	US-10-175-743-272	Sequence 272, App
24	447	36.9	283	9	US-10-176-488-272	Sequence 272, App
25	447	36.9	283	9	US-10-176-492-272	Sequence 272, App
26	447	36.9	283	9	US-10-176-747-272	Sequence 272, App
27	447	36.9	283	9	US-10-176-750-272	Sequence 272, App
28	447	36.9	283	9	US-10-176-885-272	Sequence 272, App
29	447	36.9	283	9	US-10-176-987-272	Sequence 272, App
30	447	36.9	283	9	US-10-176-991-272	Sequence 272, App
31	447	36.9	283	9	US-10-176-993-272	Sequence 272, App
32	447	36.9	283	9	US-10-184-658-272	Sequence 272, App
33	447	36.9	283	9	US-10-173-695-272	Sequence 272, App
34	447	36.9	283	9	US-10-173-697-272	Sequence 272, App
35	447	36.9	283	9	US-10-173-705-272	Sequence 272, App
36	447	36.9	283	9	US-10-174-576-272	Sequence 272, App
37	447	36.9	283	9	US-10-174-586-272	Sequence 272, App
38	447	36.9	283	9	US-10-175-747-272	Sequence 272, App
39	447	36.9	283	9	US-10-176-481-272	Sequence 272, App
40	447	36.9	283	9	US-10-176-485-272	Sequence 272, App
41	447	36.9	283	9	US-10-176-493-272	Sequence 272, App
42	447	36.9	283	9	US-10-176-756-272	Sequence 272, App
43	447	36.9	283	9	US-10-176-756-272	Sequence 272, App
44	447	36.9	283	9	US-10-176-756-272	Sequence 272, App
45	447	36.9	283	9	US-10-176-756-272	Sequence 272, App

## ALIGNMENTS

## RESULT 1

US-09-821-255-2  
; Sequence 2, Application US/09821255  
; Patent No. US20020081293A1  
; GENERAL INFORMATION:  
; APPLICANT: Michael S.C. Fung  
; APPLICANT: Bill N.C. Sun  
; APPLICANT: Cecily R.Y. Sun  
; TITLE OF INVENTION: Inhibitors of Complement Activation  
; FILE REFERENCE: 98-2A  
; CURRENT APPLICATION NUMBER: US/09/821.255  
; CURRENT FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/075,328  
; PRIOR FILING DATE: 1998-02-20  
; PRIOR APPLICATION NUMBER: 09/253,689  
; PRIOR FILING DATE: 1999-02-20  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: human  
US-09-821-255-2

Query Match 100.0%; Score 1211; DB 10; Length 228;  
Best Local Similarity 100.0%; Pred. No. 4.7e-106;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ILGREAHAHPYMASVOLNGAHLGCGVLVAEQWVLSAAHCLDEADAGKQVQLGAHSL 60		
Db	1	ILGREAHAHPYMASVOLNGAHLGCGVLVAEQWVLSAAHCLDEADAGKQVQLGAHSL 60		
QY	61	SQPEPSKRLVDLRAVPHPDSQPTIDHLLQLSEKATLGPVRLPQVDRDVPAG 120		
Db	61	SQPEPSKRLVDLRAVPHPDSQPTIDHLLQLSEKATLGPVRLPQVDRDVPAG 120		
QY	121	TLCDVAGWGVNHAGRPDSLQHVLLPVLDRATCNRRTHDGAITERLMAESNRDSC 180		
Db	121	TLCDVAGWGVNHAGRPDSLQHVLLPVLDRATCNRRTHDGAITERLMAESNRDSC 180		
QY	181	GDSGGVLVCGGVLEGGVTSGRVCGNKKPGIYTRVASYAAMIDSVLA 228		

Db 181 GDSGGLVCGGVLEGVVTSGRVCGNRKKPGIYTRVASYAAWIDSVLA 228

## RESULT 2

US-09-821-255-4

; Sequence 4, Application US/09821255

; Patent No. US20020081293A1

; GENERAL INFORMATION:

; APPLICANT: Michael S.C. Fung

; APPLICANT: Bill N.C. Sun

; APPLICANT: Cecily R.Y. Sun

; TITLE OF INVENTION: Inhibitors of Complement Activation

; FILE REFERENCE: 98-2A

; CURRENT APPLICATION NUMBER: US/09/821,255

; CURRENT FILING DATE: 2001-03-29

; PRIOR FILING DATE: 1998-02-20

; PRIOR FILING DATE: 1998-02-20

; PRIOR FILING DATE: 1999-02-20

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: FastSeq for Windows Version 4.0

; ID NO 4

; LENGTH: 233

; TYPE: PRT

; ORGANISM: pig

US-09-821-255-4

Query Match 87.2%; Score 1056; DB 10; Length 233;

Best Local Similarity 84.2%; Pred. No. 1.7e-91;

Matches 192; Conservative 18; Mismatches 18; Indels 0; Gaps 0;

QY 1 ILGREAHAHAPYMASVOLNCAHLCGGVLAQVWLSAAHCLDAADCKQVVLGAHSL 60

Db 1 ILGGEAKSHRPYMASVQVNGKHVCGGLVSEQVWLSAAHCLDAADCKQVVLGAHSL 60

QY 61 SQPEPSKRLYDLVRAVPHDPDPTIDHLLQLSEKATLGPVRLPQWRVDRDVPAG 120

Db 61 SQPEPSKRLYDLVRAVPHDPDPTIDHLLQLSEKATLGPVRLPQWRVDRDVPAG 120

QY 121 TLCVAGAGVNHAGRRPDSLOHVLPLVLDRTNCRTHDGAITERLMCAESNRDSC 180

Db 121 TLCVAGAGVNHAGRRPDSLOHVLPLVLDRTNCRTHDGAITERLMCAESNRDSC 180

QY 181 GDSGGLVCGGVLEGVVTSGRVCGNRKKPGIYTRVASYAAWIDSVLA 228

Db 181 GDSGGLVCGGVLEGVVTSGRVCGNRKKPGIYTRVASYAAWIDSVLA 228

## RESULT 3

US-09-821-255-4

; Sequence 26, Application US/10114893

; Publication No. US20020193567A1

; GENERAL INFORMATION:

; APPLICANT: Jacobs, Kenneth

; APPLICANT: McCoy, John M.

; APPLICANT: Lavallie, Edward R.

; APPLICANT: Collins-Racie, Lisa A.

; APPLICANT: Evans, Cheryl

; APPLICANT: Merberg, David

; APPLICANT: Treacy, Maurice

; APPLICANT: Bowman, Michael R.

; APPLICANT: Spaulding, Vikki

; APPLICANT: Carlin-Duckett, McKeough

; APPLICANT: Kelleher, Kerry S.

; APPLICANT: Genetics Institute, Inc.

; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

; FILE REFERENCE: GI 6000-10A

; CURRENT APPLICATION NUMBER: US/10/114,893

; CURRENT FILING DATE: 2002-04-02

; EARLIER APPLICATION NUMBER: 09/413,232

; EARLIER FILING DATE: 1999-10-06

; NUMBER OF SEQ ID NOS: 321

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 26

; LENGTH: 264

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-114-893-26

## Query Match

Best Local Similarity 42.7%; Score 456.5; DB 9; Length 264;

Matches 100; Conservative 40; Mismatches 85; Indels 9; Gaps 7;

QY 1 ILGREAHAHAPYMASVOLNCAHLCGGVLAQVWLSAAHCLDAADCKQVVLGAHSL 59

Db 27 IIGKEVSPHSRPFMASIQYGHVCGVLDIDPQWLTAAHCOYRFTKQOSTVVLGAHSL 86

QY 60 SQPEPSKRLYDLVRAVPHDPDPTIDHLLQLSEKATLGPVRLPQWRVDRDVPAG 119

Db 87 LSKNEASKOTLEIKKFIPIPSRVTSPOQSNIDIMVLKLOTAALKNHVKML-HIRSKTSLS 145

QY 120 GTLCDVAGVNHAGRRP-DSLOHVLPLVLDRTNCRTHDGAITERLMCA-ESN 174

Db 146 CTCKCVTGWATDPSLRPSDTLREVTVLSRKLCSQSYNGDPFITKDMVCAGDAGK 205

QY 175 RRDCKGDSGGPLVCGGVLEGVVTSGRVCGNRKKPGIYTRVASYAAWIDSVL 227

Db 206 QKDSCKGDSGGPLVCGGVLEGVVTSGRVCGNRKKPGIYTRVASYAAWIDSVL 258

## RESULT 4

US-09-888-615-112

; Sequence 112, Application US/09888615

; Patent No. US20020064856A1

; GENERAL INFORMATION:

; APPLICANT: PLOWMAN, GREGORY

; APPLICANT: WHYTE, DAVID

; APPLICANT: CAENEPEEL, SEAN

; APPLICANT: CHARYDCZAK, GLEN

; APPLICANT: MANNING, GERARD

; APPLICANT: SUDARSANAM, SUCHA

; TITLE OF INVENTION: NOVEL PROTEASES

; FILE REFERENCE: 038602/1214

; CURRENT APPLICATION NUMBER: US/09/888,615

; CURRENT FILING DATE: 2001-06-26

; PRIOR APPLICATION NUMBER: 60/214,047

; PRIOR FILING DATE: 2000-06-26

; NUMBER OF SEQ ID NOS: 150

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 112

; LENGTH: 284

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-888-615-112

Query Match 37.2%; Score 451; DB 10; Length 284;

Best Local Similarity 42.7%; Pred. No. 1.3e-34;

Matches 100; Conservative 34; Mismatches 88; Indels 12; Gaps 6;

QY 1 ILGREAHAHAPYMASVOLNCAHLCGGVLAQVWLSAAHCL-EDAADCKQVVLGAHSL 58

Db 35 IIGKEVTPHSRPFMASYVFGGQHHCGGFLRLARVWVSAAHCFSHRDLRTG-LVVLGAH 92

QY 59 SLQPEPSKRLYDLVRAVPHDPDPTIDHLLQLSEKATLGPV-RPLPQWRVDRD 116

Db 93 VLSTAEPTQOVFGIDALTTDPYHPMTHANDICLLQLNGSAVLGPVGLLRUPGRR-RP 151

QY 117 VAPGTLCDVAGVNHAGRRPDSLOHVLPLVLDRTNCRTHDGAITERLMCA-ESN 173

Db 152 PTAGTRCRVAGNGFVSDFEELPPGLMEAKVRVLDPDVCN--SSNKGHLTLMLCTRSGDS 209

QY 174 NRRDCKGDSGGPLVCGGVLEGVVTSGRVCGNRKKPGIYTRVASYAAWIDSVL 227

Db 210 HRRGFCSDGSGPLVCGGVLEGVVTSGRVCGNRKKPGIYTRVASYAAWIDSVL 263